Der Einfluß von Prüfungssituationen auf das Leistungsverhalten von Schülern im 4. und 9. Schuljahr

Inaugural - Dissertation
zur Erlangung des Doktorgrades
der

Philosophischen Fakultät

der
Westfälischen Wilhelms-Universität zu
Münster (Westf.)

1973

vorgelegt von
Friedhelm Holfort
aus Essen

1972

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547202_0032

Psychologie

Der Einfluß von Prüfungssituationen auf das Leistungsverhalten von Schülern im 4. und 9. Schuljahr

Inaugural - Dissertation
zur Erlangung des Doktorgrades
der

P h i l o s o p h i s c h e n F a k u l t ä t

der
Westfälischen Wilhelms-Universität zu
Münster (Westf.)

vorgelegt von
Friedhelm Holfort
aus Essen

1 9 7 2

UNIV.
LEIDEN
BIBL.

Tag der mündlichen Prüfung: 22. Januar 1973

Dekan: Prof. Dr. Herbert Mainusch

Referent: Prof. Dr. Dr. h.c. Wolfgang Metzger

Korreferent: Prof. Dr. Ernst Bornemann

INHALT

Seite

1. EINFÜHRUNG IN DEN THEMENBEREICH, FRAGESTELLUNG UND STANDORTBESTIMMUNG DER UNTERSUCHUNG 3
1.1. Schule und Leistung in der heutigen Industriegesellschaft 3
1.2. Fragestellung 4
1.3. Standort der Untersuchung 7

2. ANLAGE DES EXPERIMENTS..... 9
2.1. Methodologische Vorbemerkungen - Zur Dynamik und Struktur von Leistungsprozessen 9
2.2. Forschungsansatz 15
2.3. Vorversuche und Vorarbeiten 17
2.4. Auswahl der Versuchspersonen 19
2.5. Analyse des Versuchsmaterials 20
2.5.1. Diktate 22
2.5.2. Zahlenrechnen 24
2.5.3. Analyse der Fragebogen 27
2.6. Anordnung und Durchführung der Versuche 29
2.7. Anweisungen (Beispiel) 34

3. EINFLUß DER PRÜFUNGSBEDINGUNGEN 37
3.1. Leistungsunterschiede im Gesamtergebnis beider Schuljahre 38
3.2. Unterschiede im Leistungsverlauf 46

4. EINFLUß VON PERSÖNLICHKEITSMERKMALEN 56
4.1. Prüfungssituation und Fachleistung 57
4.2. Prüfungssituation und Geschlecht 62
4.3. Prüfungssituation und Einstellung zum Schulfach 65
4.4. Prüfungssituation und Zeichensetzung 66

Seite

5. ANSÄTZE EINER THEORIEBILDUNG 69

6. ZUSAMMENFASSUNG 76

7. DISKUSSION DER ERGEBNISSE UND
DES ANSATZES 81
7.1. Auseinandersetzung mit der Literatur 81
7.2. Pädagogische Erörterungen: Reduzierung
von Angst- und Belastungszuständen
im Schulalltag 93

8. AUFGABEN FÜR DIE WEITERE FORSCHUNG 98

LITERATUR 103

ANHANG

1. EINFÜHRUNG IN DEN THEMENBEREICH, FRAGESTELLUNG UND STANDORTBESTIMMUNG DER UNTERSUCHUNG

1.1. Schule und Leistung in der heutigen Industriegesellschaft

In den letzten Jahrzehnten ist darauf hingewiesen worden, daß der Schulbetrieb in seiner Gesamtheit - sieht man einmal von den Waldorf-, Montessori-Schulen und Zweigen mit ähnlichen pädagogischen Intentionen ab - hinsichtlich seiner Leistungsforderungen ein Abbild der modernen Industriegesellschaft ist. Diese Gesellschaft honoriert Konformität des Individuums bezüglich der erwarteten Leistungsbereitschaft, -anpassung und -fähigkeit mit der Befriedigung weitgehender individueller Bedürfnisse (materielle Entschädigung, Gruppenzugehörigkeit, angesehene Stellung in der Gruppe). Auch in der Schule werden Fähigkeiten, Kenntnisse und Fertigkeiten verlangt und geprüft, wie es sehr deutlich in den Richtlinien und Lehrplänen des Bundeslandes NRW für die Grund- und Hauptschulen niedergelegt ist (Richtlinien für die Grundschule 1969, S. 156/157/455 ff; Richtlinien für die Hauptschule 1968, B 3/13, B 355, B 7/5). Die Leistungsfeststellungen nehmen nach wie vor einen bedeutsamen Bereich im Unterrichtsgeschehen der modernen Schule ein, zumal die schulischen Leistungsnachweise des Individuums das Niveau und die Ausrichtung der weiteren Ausbildung entscheidend mitbestimmen. Die existentielle Bedeutung der Qualität der Leistungsnachweise für das Leben der Schüler kann durch Feststellungen von SCHELSKY nur unterstrichen werden. SCHELSKY (1959) sieht "die Gesellschaft unserer Zeit als nivellierte Berufsgesellschaft mit starkem familienindividuellen Aufstiegsstreben an. Die erhöhten theoretischen Anforderungen in vielen Berufen der modernen Industriewelt machen

den Aufstieg in leitende Stellungen in vorher nicht gekanntem Maße vom Nachweis höherer Schulbildung, vom Berechtigungsschein abhängig. Das Sozialprestige der Familien, das Streben nach höherem Einkommen und sozialer Sicherheit scheinen von diesem Nachweis abhängig zu sein. Die Schule wird zur entscheidenden zentralen sozialen Dirigierungsstelle, zur bürokratischen Zuteilungsapparatur von Lebenschancen" (Schule als Steuerungsinstanz der sozialen Mobilität - SCHELSKY 1962). In ähnlichem Sinne äußert sich HECKHAUSEN (1968). Auch eine ganztägige Gesamtschule wird nach abgeschlossener Entwicklung den determinierenden Einfluß der schulischen Leistungsnachweise für die Einordnung in das Gesellschaftssystem nicht einschränken, da eindeutige Leistungskriterien über die Zulassung zu qualifizierten Leistungsgruppen entscheiden werden, wie es auch SALZMANN (1966) und VON HENTIG (1971) vermuten.

1.2. Fragestellung

Derzeit erfüllen die Grundschule wie auch die weiterführenden Schulen die gesellschaftliche Auslesefunktion der Grobsortierung, indem sie mit den wenig zuverlässigen Methoden des mündlichen Abfragens, der Klassenarbeit und des Zeugnisses (hier besonders des Abschlußzeugnisses) das Erreichen bestimmter Lehrplanziele prüfen. Die Betonung dieser 'Zielleistungen' für das Erreichen des Klassenziels oder - über einen längeren Zeitraum hinweg geplant - Erringen von angesehenen Berufspositionen hat es mit sich gebracht, daß Leistungsfeststellungen der Schule - insbesondere Klassenarbeiten - zu herausragenden Marksteinen für Schüler und Eltern im Schulalltag werden. Die Analyse solcher Prüfungssituationen wirft neben anderen reizvollen Varianten zwei Problemstellungen auf, die sich als frucht-

bare Ansätze für wissenschaftliche Forschungen erweisen können.

a. Quantitative Veränderungen in der Gesamtleistung bzw. im Leistungsverlauf während schulischer Prüfungssituationen

b. Ablauf und Struktur der dynamischen Prozesse in schulischen Prüfungssituationen

Beide Themenkomplexe können zusammen nur im Rahmen eines größeren Forschungsprogrammes methodisch sauber, umfassend und sorgfältig analysiert werden. Der Verfasser entschied sich auf Grund interessanter Hinweise von SANDVOSS (1952, 1956), wertvoller Anregungen von METZGER (1957) und einiger von ihm selbst während seiner schulpraktischen Ausbildung durchgeführter Erkundungsexperimente (HOLFORT 1969) zur Bearbeitung der unter a. verzeichneten Thematik, indem er schon erprobte Ansätze (SANDVOSS 1952, HAECKER 1964, MC COY 1965) weiterführte und ausbaute. In diesen Untersuchungen hatte es sich gezeigt, daß Informationen und Instruktionen den belastenden Charakter von Probesituationen ändern können. Zu diesen stimulierenden Informationen dürfen auch die Hinweise auf die schulischen Zwischen- oder Endziele gezählt werden, die ein Schüler erreichen möchte.

In der normalen schulüblichen Übungsarbeit werden alle Hinweise auf Leistungskontrollen unterlassen, um die Situation möglichst frei von Leistungsangst zu halten, die nach Meinung vieler Autoren (vgl. Kap. 1.3.) unmittelbar mit dem Phänomen Prüfungssituation zusammenhängt. Die gewohnte Klassenarbeit vollzieht sich zwar schon als Prüfungsritual, ist aber nicht letztentscheidendes, einmaliges Bewertungsindiz für ein Endziel. Ein Teil der Belastung wird durch das Wissen genommen, daß man noch Chancen zur Verbesserung hat. Die Prüfungsarbeit wird den

Schülern per Instruktion und Ritual als entscheidendes Leistungsindiz mit irreparablen Folgen für die Note im Abschlußzeugnis vorgestellt. Die Bedeutung der Schulzeugnisse beim Übergang zu weiterführenden Schulen oder zu angestrebten Lehrstellen erhöht zudem das Gewicht der Probearbeiten in diesen Schuljahren.

Aus diesen Überlegungen ergibt sich folgende Fragestellung:

> Wie unterscheiden sich die Leistungen von Schülern des 4. und 9. Schuljahres, wenn man ihre Rechen- und Rechtschreibleistungen in drei schulnahen Prüfungssituationen mit unterschiedlich anregenden Instruktionen und Informationen untersucht? Wie vollzieht sich das Zusammenspiel von experimentellen Bedingungen und bestimmten Persönlichkeitsvariablen?

Zunächst wird der Standort der Untersuchung innerhalb des großen Forschungsbereichs: Angst und Leistung bestimmt. Nach einer Analyse der methodologischen Grundlagen dieses Feldexperiments und der Befunddarstellung wird versucht, erste Ansätze einer Theoriebildung hinsichtlich des Problems der Leistung in schulischen Prüfungssituationen zu diskutieren. Schließlich werden die Ergebnisse unter Einbeziehung der in der Literatur dargestellten Hypothesen erörtert, Möglichkeiten der Angstreduzierung im Schulalltag vorgestellt und Aufgaben für zukünftige Feldexperimente besprochen.

Für die Leitung und Betreuung dieses Unterrichtsversuches bin ich Herrn Prof. Dr. Dr. h.c. WOLFGANG METZGER zu besonderem Dank verpflichtet. Dank gebührt auch den Kolleginnen und Kollegen vieler Schulen, die ihre Klassen bereitwillig für diesen

Unterrichtsversuch zur Verfügung stellten. Die Fertigstellung der Arbeit verdanke ich nicht zuletzt der Mithilfe meiner Frau bei der Zweitkorrektur der Prüfarbeiten.

1.3. Standort der Untersuchung - Überblick: Forschungsansätze und Beiträge zum Problemkreis Leistung und Prüfungssituation (Angst)

Inhaltlich kann der vorliegende Ansatz als Beitrag zum allgemeinen Thema Leistung und Angst gerechnet werden, da Angst unlösbar mit dem Phänomen der Prüfung verknüpft zu sein scheint (METZGER 1971, KVALE 1972).

Im angelsächsischen Sprachraum entwickelte sich ein besonderer Zweig der Psychologie, der reichhaltige empirische Befunde im Rahmen dieser Problematik erbrachte und die bis dahin kaum experimentell orientierte Pädagogik zu einer fruchtbaren Überprüfung ihrer wohl durchdachten, aber empirisch nicht gesicherten Grundsätze anregte.

Eine Analyse der bis heute vorliegenden Experimente legt allerdings den Schluß nahe, daß noch kein allgemein akzeptiertes Konzept der Angst in Leistungssituationen erarbeitet wurde (LEVITT 1971, GÄRTNER-HARNACH 1972).

Theoretische Ansätze wie die Iowa-Theorie (drive-Modell) von TAYLOR (1956) und SPENCE (1958), die Yale-Theorie (habit-Interferenz-Modell) von MANDLER und SARASON (1952), das Erwartungsmodell der Angst von ATKINSON und LITWIN (1960) und - in jüngster Zeit - das Modell von THURNER (1970), wie auch eklektische Bemühungen von EYSENCK (1957) erweisen sich im praxisnah angelegten Feldexperiment als nicht so umfassend, als daß sie ein übergreifen-

des, gesichertes Modell zur Erfassung und zum Verständnis aller Wirkvariablen bilden können. Die Komplexität des Faktorengefüges und die Situationsgebundenheit aller Erkenntnisse im realen psychischen Gesamtfeld behindern die Entwicklung eines allgemein gültigen Konzepts. Angesichts der Vielzahl unterschiedlicher Definitionen und Konzepte der Angst (WESTERMAN HOLSTIJN 1966) beschränkt sich die vorliegende Arbeit auf Feststellung quantitativer Leistungsmerkmale in festgelegten realen schulischen Prüfungssituationen.

Konkrete experimentelle Ansätze zu dieser Thematik liegen nur in geringem Umfange vor. Man kann Untersuchungen ausklammern, die das Prüfungsverhalten von Studenten erkunden (MOELLER 1969, HELM 1954). Ergebnisse aus Experimenten, die Beziehungen zwischen Angstmaßen (Fragebogen und Tests zur Messung der Leistungsangst: TAQ, TASC, AAT u.a.) und Intelligenzwerten herausstellen, können das Faktorengefüge im Problemzusammenhang Leistung in schulischen Prüfungssituationen nur wenig aufhellen (MANDLER und COWEN 1958, DURNFORD 1957, BROEN 1959, FELDHUSEN und KLAUSMEIER 1962, BIGGS 1962, SILVERSTEIN 1966). Auch Studien über den Einfluß von Leistungsangst auf schulische Erfolge und die Fähigkeit zum Lösen von Problemen können nur in Teilergebnissen für die Interpretation der Ergebnisse fruchtbar sein, da auch sie spezifische Variable der schulischen Prüfungssituation außer acht lassen (DORIS 1959, ALPERT und HABER 1960, LJUNG 1960, COX 1962, BIGGS 1962, GIFFORD und MARSTON 1966).

Im deutschsprachigen Raum haben nur wenige experimentelle Analysen Gewicht, sieht man von den zahlreichen weniger bedeutungsvollen nichtexperimentellen Fallstudien und Erfahrungsberichten ab (HANSELMANN 1950, TRAMER 1952, MIERKE 1953, WIDMER 1955, FRUHMANN 1955, HOCHHEIMER 1955, GRAAP 1956, ECKSTEIN 1958, LENNERT 1959, CORRELL 1962,

GERÉB und SZABO 1963). Zu den experimentellen Arbeiten gehört die Untersuchung von SANDVOSS (1952), der die Leistungen von Zweitkläßlern im schriftlichen Zahlenrechnen in leistungsdruckarmer Übungssituation mit Arbeitsprodukten der Klassenarbeit vergleicht. Erwähnenswert ist auch eine Arbeit von HAECKER (1964), der einen SANDVOSS nachempfundenen Unterrichtsversuch in einer Volksschule (1. - 7. Schuljahr) durchführt. In diesen Kontext gehören Befunde von BARTMANN (1962), der den Einfluß von Zeitdruck untersucht, weiterhin Ergebnisse von ZIELINSKI (1967), der Beziehungen zwischen Ängstlichkeit, schulischer Aktivität, Intelligenz und Schulleistung bei neun- bis elfjährigen Volksschülern herausstellt. TAUSCH und Mitarbeiter (1969) prüfen die Möglichkeiten der Verminderung von Angst und Neurotizismus durch positive Bekräftigungen. Die erwähnten Untersuchungen sollen den Problemzusammenhang, in dem die vorliegende Arbeit eingebettet ist, aufhellen.

Zahlreiche Untersuchungen aus dem angelsächsischen Raum, die für die Beantwortung der Fragestellung nur wenig Relevanz besitzen, fehlen in dieser Zusammenstellung. Vollständigere, umfassende Darstellungen des Schriftenverzeichnisses zum Problemkreis: Leistung und Angst finden sich bei SÜLLWOLD (1959) im Anschluß an sein Sammelreferat, GÄRTNER-HARNACH (1972) und SCHELL (1972).

2. ANLAGE DES EXPERIMENTS

2.1. Methodologische Vorbemerkungen - Zur Dynamik und Struktur von Leistungsprozessen im pädagogischen Feld

Wie alle Leistungsprozesse sind auch diejenigen im

schulischen Bereich, im komplexen pädagogischen Feld, nicht aus dem strukturellen und dynamischen Zusammenhang mit realen situativen Momenten der Schulwirklichkeit herauszulösen. BECK (1953) betont die "ganzheitlich-dynamische Vollzugseinheit zwischen leistender Person und der zu bewältigenden Situation" und schließt damit an LEWIN und seine Mitarbeiter an, die nach Formulierung des dynamischen Ansatzes der Feldtheorie aus der Kontroverse mit dem statischen Assoziationismus ACHS heraus dazu übergingen, komplexe soziale Situationen zu untersuchen (LEWIN und Mitarbeiter ab 1926). Sie weisen nach, daß innerhalb der "ganzheitlich-dynamischen Vollzugseinheiten" einzelne Zustände und Vorgänge einander bedingen und tragen. Jede Änderung eines Elements hat Wirkungen, die einen Umbau des Ganzen bedeuten können und sich auch in der Formel LEWINS zur Erklärung des psychischen Geschehens (Verhaltens) widerspiegeln:

$$V = f\,(P, U).$$

Psychisches Geschehen (Verhalten, Leistung) ist eine Funktion der Person (P) und ihrer Umgebung (U) (nach LEWIN 1936, 1969).

In ähnlichem Sinne betont WERTHEIMER (1945, 1957) in seinem für Erzieher bedeutsamen Werk: Produktives Denken: "Aber wenn wir wirklich verstehen wollen, wie die Tätigkeit (ein mathematisches Problem zu lösen) zustande kommt (oder auch nicht zustande kommt), müssen wir ein viel breiteres Feld ins Auge fassen. Die Frage richtet sich dann auf die Organisation des gesamten Feldes, innerhalb dessen das ganze augenblickliche Geschehen selbst nur ein Teil ist - des persönlichen, des sozialen, des historischen Feldes" (S. 75). Grundlegende Untersuchungen LEWINS und seines Kreises wie auch andere

wegweisende Arbeiten aus dem Bereich der Gestaltpsychologen (KÖHLER 1940, 1947, KOFFKA 1935, WERTHEIMER 1945, 1957, METZGER 1954, 1962) trugen dazu bei, "angemessene wissenschaftliche Werkzeuge für die Untersuchung von Gestalten und Strukturen im psychischen Gesamtfeld zu entwickeln" (WERTHEIMER 1945, 1957).

So kann die vorliegende Fragestellung auch nur dann hinreichend beantwortet werden, wenn quantitative Leistungsänderungen im dynamischen Zusammenhang mit anderen realen Variablen der Schulwirklichkeit studiert werden (vgl. auch die später erwähnten Forderungen von WINNEFELD 1957).

Ein Beweis für die Bedeutung und Gültigkeit des dynamischen feldtheoretischen Forschungsansatzes LEWINS ist der weitgesteckte Rahmen, innerhalb dessen es zu einer "fruchtbaren Verbindung zwischen Pädagogischer, Angewandter und Sozialpsychologie, die ihren Schwerpunkt in der dynamisch-topologisch dargestellten Sozialforschung hat, kommen kann" (SOLLE 1969).

E. KÖHLER, PETERSEN und Mitarbeiter regten Anfang der dreißiger Jahre die Erforschung der realen Unterrichtssituation mit Hilfe ausführlicher Protokolle des Lehrer- und Schülerverhaltens an. WINNEFELD, nach SOLLE (1969) ein "maßgeblicher Vertreter einer pädagogischen Feldtheorie", bemühte sich im Sinne seiner Vorgänger und im Anschluß an die von LEWIN entwickelte Feldtheorie um eine dementsprechend angemessene, methodologisch durchdachte Analyse typischer Unterrichtssituationen und Strukturen von Geschehensabläufen mit der Absicht, die Ausläufer einer rein spekulativ und normativ orientierten Pädagogik endgültig zu überwinden und eine

empirische Grundlage zu sichern. Dieses Unternehmen stieß auf Schwierigkeiten, die hauptsächlich durch den unterrichtlichen Geschehensablauf bedingt sind: Personenzahl, Organisationsprobleme, Unterrichtsmethoden, Überdeterminiertheit vieler Verhaltensweisen, historische Abläufe (Kindheitserlebnisse), Milieueinflüsse.

An Hand der Aussagen WINNEFELDs sollen die Feldstrukturen der pädagogischen Situation verdeutlicht werden.

WINNEFELD[1)] begreift "pädagogische Geschehensbereiche als sozialpsychische Spannungsfelder hochkomplizierter Art", die sich durch "Störbarkeit, Offenheit und Labilität" auszeichnen. Diese Einschätzung des pädagogischen Geschehensablaufs deutet Schwierigkeiten der Unterrichtsforschung an, die vor allem die "Intimität der pädagogischen Situation" nicht zerstören darf. Dies glaubt WINNEFELD durch Einschaltung einer längeren Gewöhnungsperiode vor der eigentlichen Untersuchung gewährleisten zu können. Das spezifisch Pädagogische ist die "bestimmte teleologische Strukturiertheit", die in einem "Gefällefeld" besteht, "in dem zwischen den darin befindlichen Menschen ein bestimmtes Reifegefälle vorhanden ist", das auf Ausgleich drängt, in das aber auch situative "gesellschaftliche und geschichtliche Faktoren" hineinragen und sowohl den Geschehensablauf als auch die Interpretation von Forschungsergebnissen in diesem Feld beeinflussen. Empirische Befunde bleiben "gleichsam situationsverhaftet" ("historisch - teleologisch - situative Gebundenheit" der Erkenntnisse).

[1)] Die folgenden Ausführungen beziehen sich auf F. WINNEFELD, Päd. Kontakt und pädagogisches Feld, Schule und Psychologie, Nr. 7 Basel 1957, 30 - 44

Jedes pädagogische Feld zeichnet sich durch eine "ungeheure Vielfalt mannigfaltig verflochtener und ineinander übergreifender, sich gegenseitig bedingender Faktoren aus". Dem Menschen, der sich in diesem Feld bewegt, erkennt WINNEFELD einen "Spontaneitätsspielraum" zu, der die Anerkennung eines eindeutigen Reiz - Reaktions - Mechanismus nicht zuläßt, stattdessen vielfältige Antwortmöglichkeiten auf äußerlich gleiche Reizsituationen erwarten läßt, wodurch das Vorhersagen des individuellen Verhaltens erschwert wird. Jeder pädagogische Geschehensverlauf, den WINNEFELD als "komplexe Handlungsgestalt" auffaßt, enthält Teile (Unterganze), deren Stellenwert nur in der Betrachtung der Gesamtlage gewonnen werden kann. Dies bedeutet für die Interpretation von Untersuchungsergebnissen im pädagogischen Feld: Sowohl Einzelergebnisse als auch Gesamtergebnisse gewinnen erst Gewicht in der Berücksichtigung ihrer historisch - teleologisch - situativen Gebundenheit, des spezifischen Zusammenhangs ihrer Gliederfunktion mit dem gesamten Geschehensverlauf. WINNEFELD deutet an, daß es sich nicht um die "Allerfassung", also eine "Gesamtschau" einer Situation handelt, sondern um eine "Ganzerfassung" der pädagogischen Situation, "indem sie zwar Einzelfragen angeht, bei der Einordnung und Wertung von Einzelbefunden jedoch stets den Blick auf das beobachtete aktuelle pädagogische Ganzgefüge in seiner Komplexität zurücklenkt." Konkrete Untersuchungen zur Struktur und Dynamik von Problemlösungsprozessen von WERTHEIMER (1945, 1957), DUNCKER (1935) und METZGER (1962) unterstreichen die Fruchtbarkeit und Gültigkeit dieses Forschungsansatzes.

Als letztes und für psychologische Analysen des pädagogischen Geschehensablaufes besonders relevantes Struktur-

merkmal stellt WINNEFELD die Gesetzmäßigkeit des Geschehensverlaufes heraus, die erst das Herausarbeiten "typischer Bedingungsstrukturen" sichert. Obwohl jeder pädagogische Geschehensverlauf in sich einmalig und unwiederholbar ist, enthält er meist "gesetzliche Bezüge" (vgl. auch METZGER 1954), die als "Beziehungssystem" erfaßt werden.

Die vier Hauptstrukturzüge pädagogischer Felder sind:

Labilität, Offenheit, Störbarkeit,
historisch - teleologisch - situative Gebundenheit,
Vieldimensionalität des Faktorengefüges,
Gesetzmäßigkeit des pädagogischen Geschehensverlaufs.

Sie erfordern nach WINNEFELD im Rahmen empirischer Untersuchungen sowohl "resolutive" als auch "kompositive" Schritte: "Zuerst hat eine Aufgliederung des Wirklichen zu erfolgen, wobei Beziehungen isolierend erkannt werden. Es müssen die verschiedenen Relationsordnungen entdeckt werden. Danach gilt es, die aufgefundenen Beziehungen zusammenzufügen. Die erhaltenen Teilsysteme sind zu einem vollständigen Bild des Gesamtvorganges zu verknüpfen, die aber nicht mehr in ungegliederter Totalschau, sondern als differenziertes Beziehungsganzes" angesehen werden.

Dieses methodische Konzept gilt nicht nur für die von WINNEFELD angestrebten Beobachtungsstudien im pädagogischen Feld, sondern auch für die anderen im natürlichen Milieu ablaufenden Experimente (auch Unterrichtsversuche) (PAGÈS 1962, Über Feldexperimente bei FRENCH 1953, 1966).

2.2. Forschungsansatz dieser Untersuchung

Der vorliegende Ansatz gehört zum Bereich von wissenschaftlichen Untersuchungen, die sich mit der Auswirkung von Prüfungssituationen in der Schulwirklichkeit befassen. Prüfungssituationen können mit Stress-Situationen (vgl. SANDVOSS 1952, METZGER 1962) identifiziert werden, wenn man sich Auffassungen von SELYE (1957) und JANKE (1969) anschließt und diese mit den Wirkungen der schulischen Prüfungssituation vergleicht. Als Stressoren werden nach JANKE sämtliche auf den Organismus wirkende Stimuli angesehen, sobald sie eine gewisse Intensität und/oder Dauer überschreiten. Alle Stressoren zeichnen sich durch folgende Gemeinsamkeiten aus:

a. Sie erzeugen die Auslösung von negativen Affekten (Angst, Furcht, Unlust, Ärger),
b. sie fördern psychologische Verarbeitungs- und Adaptionsvorgänge, die auf eine Reduktion dieser negativen Affekte und eine Überwindung der Stress-Situationen abzielen.

Ziel dieser Anpassungserscheinungen ist (auch nach SELYE 1957) die Homöostase, die Überführung des Organismus in den normalen Ruhezustand. Von den wichtigen Gruppen von Stressoren: schmerzerregende Stressoren, überhöhte sensorische Reizung, gedrosselte sensorische Reizung, soziale Isolation, sei in diesem Zusammenhang besonders auf die frustrierenden Situationen verwiesen.

LAZARUS et al. (1952) entwickelten für die experimentelle Erforschung solcher Stress-Situationen zwei wesentliche Möglichkeiten (zit. nach BARTMANN 1963):

a. "Der Proband wird seelischem Druck ausgesetzt. Man läßt ihn Mißerfolge bei der Arbeit an unlösbaren oder während der Durchführung unter-

brochenen Aufgaben erleben oder man erhöht seine Furcht vor Mißerfolg durch Angabe falscher Vergleichszahlen für seinen Begabungs- und Leistungsstand.

b. Es wird der Einfluß einer gesteigerten Beanspruchung auf das Verhalten der Probanden untersucht. Man gestaltet die Arbeitssituation so, daß sie außerordentliche Anforderungen an den Probanden stellt (Mehrfacharbeit, Zeitkontrolle, zusätzliche akustische und optische Signale."

Diese Möglichkeiten betreffen weniger praxisnahe Untersuchungen, auch Laboratoriumsexperimente, die das oft vielschichtige, voneinander abhängende, komplexe Beziehungsgefüge dynamischer psychischer Prozesse in der realen Wirklichkeit oft nur unvollständig beschreiben können.

Eine dritte Möglichkeit soll noch erörtert werden, um den methodischen Ansatz dieser Untersuchung zu lokalisieren.

c. Hier werden in der schulischen Normalsituation vor allem quantitative Phänomene der Auswirkung der seelischen Belastung untersucht. Methodische Mängel werden dadurch ausgeschaltet, daß man schulische Prüfungssituationen so weit wie möglich standardisiert, ohne systemändernde Eingriffe vorzunehmen, und eine möglichst große Anzahl von Schülern berücksichtigt.

Die Analyse solcher Stress-Situationen des Schulalltags ist bisher nur selten ins Auge gefaßt worden, weil schwer lösbare Probleme der Standardisierung von Verfahrensweisen und mangelnde Kontrolle äußerer Faktoren die Qualität empirischer Befunde beeinflussen können.

2.3. Vorversuche und Vorarbeiten

Im Sommer 1969 sowie im Sommer 1970 führte der Verfasser zahlreiche Erkundungsexperimente an Haupt- und Grundschulen im Bereich des Schulamtes Essen (vor allem im Schulaufsichtsbezirk II) zur Abklärung des methodischen Vorgehens durch. Anläßlich der schriftlichen Arbeit zur zweiten Staatsprüfung für das Lehramt an Volksschulen wurde der Problemkreis 'Zahlenrechnen' in neun Leistungsgruppen an einer Hauptschule (A, B, C - Gruppierung) untersucht. Ähnliche Vorversuche im 4. Schuljahr der Grundschule ergaben aufschlußreiche Hinweise für die Planung des Diktierversuchs. Die Serie der Erkundungsexperimente wurde ergänzt durch Beratung und Diskussion mit erfahrenen Schulpraktikern und Spezialisten im Hinblick auf die anstehenden fachlichen und didaktischen Probleme.

Als wegweisend für die gesamte weitere Untersuchung erwiesen sich folgende Erkenntnisse:

a. Die Mehrzahl der Lehrer verhielt sich nur unter Kontrolle gemäß den präzise formulierten Versuchsanweisungen. Ein solches Feldexperiment in der Schule konnte unter der Prämisse: Identität der Versuchsbedingungen und Vergleichbarkeit der Ergebnisse nur dann zufriedenstellend durchgeführt werden, wenn das Lehrerverhalten ausreichend kontrolliert wurde oder ein oder mehrere vorher geschulte Versuchsleiter die Versuche durchführten.

b. Schülergruppen, die z.B. an Versuchsschulen unter nicht üblichen Bedingungen (vornehmlich Gruppenarbeit) unterrichtet wurden, Klassen mit unverhältnismäßig hohem Anteil an Gastarbeiterkindern (über 20 %), besonders kleine oder große Gruppen

(weniger als 12, mehr als 45 Schüler) wurden ausgeschieden.

c. Das struktur- und schwierigkeitsidentische Versuchsmaterial, wie z.B. das 'Päckchenrechnen' (vgl. auch SANDVOSS, HAECKER), erwies sich einmal als wenig schulnah, zum anderen als hemmend, d.h. wenig belebend für den Arbeitseifer der Versuchspersonen.

d. Weder neu eingeführte Hefte noch die angebotenen DIN-A-4-Blätter konnten den angestrebten Alltagscharakter der Übungsarbeiten (Schülerverdacht: Lehrerkontrolle) und die Ernstsituation der Klassenarbeit (Schülerargument: Bedeutungslose Arbeit außerhalb der Reihe) gewährleisten. Deshalb benutzten die Schüler im Hauptversuch die gewohnten Übungs- oder 'Schmierhefte' und die vertrauten Klassenarbeitshefte.

e. Im Zahlenrechnen wurde die Anzahl der Aufgaben gefunden, die vom Durchschnitt der Schüler in etwa 25 - 30 Minuten bearbeitet sein konnte (12 Aufgaben für das 4. Schuljahr, 15 Aufgaben für das 9. Schuljahr). Das Aufgabenmaterial mußte so beschaffen sein, daß es einmal differenzierte Aussagen bezüglich der Leistung unter variierten Prüfungsbedingungen zuließ, zum anderen aber auch Verminderung der Leistung durch überlange Arbeitsdauer ausschaltete und verhinderte, daß Zeitdruck und das dadurch bedingte Anwachsen der Motivationsspannung die Leistung beeinflußte. Seit WAGENSCHEIN (1958) und BARTMANN (1962) kennt man den Einfluß des Zeitdrucks auf spezifische Arbeitsleistungen bei vielen Schulkindern. Deshalb wurden alle Versuchspersonen vor den Versuchen davon unter-

richtet, daß das Klingeln der Schulglocke nicht das Ende der Arbeitszeit anzeigte.

Die Diktattexte wurden auf Schwierigkeit (Bestätigung der von GLATZ angegebenen Fehlerquoten), Umfang und Inhalt geprüft.

Das Verhalten des Lehrers wurde in kleine Einheiten zerlegt und schriftlich festgehalten. So konnte ein Kanon von Verhaltensweisen erarbeitet werden, der das Unterrichtsgebaren des Versuchsleiters weitgehend standardisierte, wodurch praxisfremde, somit verdachterregende Eingriffe, die den Experimentiercharakter der Versuchsserie offengelegt hätten, vermieden wurden und fast identische Versuchsbedingungen in allen Klassen gesichert werden konnten.

Versuchsanweisungen wurden in zahlreichen Varianten auf ihren Effekt und ihre Nützlichkeit hin überprüft, verworfen oder endgültig festgelegt.

2.4. Auswahl der Versuchspersonen

Die Versuchsserien sowie alle Kontrollserien wurden unter der Leitung des Verfassers in der Zeit von September 1970 bis Ende September 1971 durchgeführt. Rund 700 Schülerinnen und Schüler aus den Abschlußklassen von Essener Hauptschulen sowie rund 800 Kinder aus vierten Schuljahren der Grundschulen aus verschiedenen Aufsichtsbezirken der Schulämter Essen, Duisburg, Oberhausen, Moers und Emmerich nahmen daran teil. Die Schulen wurden nach Eingang der Durchführungsgenehmigung des Kultusministeriums von NRW und Absprache mit Schulräten, Rektoren und Lehrern so ausgewählt, daß die Schülerzusammensetzung in etwa die soziologische Bevölkerungsstruktur des Ruhrgebiets repräsentierte.

Die Untersuchung im Zahlenrechnen (Kontrolle der Übungsarbeit) war so angelegt, daß nur je eine Klasse oder Leistungsgruppe pro Schule teilnehmen konnte. Im Zahlenrechnen (9. Schuljahr) wurden nur B-Leistungsgruppen oder undifferenzierte Klassenverbände zugelassen, die wenigstens ein Jahr oder länger als einheitliche Gruppe unterrichtet wurden, weil sich für A-Leistungsgruppen das angebotene Aufgabenmaterial in Vorversuchen als zu leicht (zu geringe Durchschnittsfehlerzahl) erwies und in C-Gruppen oftmals nicht die angestrebte Identität der Versuchsbedingungen zu erreichen war (mangelnde Leistungsbereitschaft, Resignation, Aggressivität der Schüler, ungenügende Vorkenntnisse, oft mangelhafte pädagogische und fachliche Qualifikation des Fachlehrers).

2.5. Analyse des Versuchsmaterials

Das Versuchsmaterial mußte so beschaffen sein, daß es sich einmal inhaltlich wie formal kaum vom Üblichen Schulstoff abhob, zum anderen vergleichbare Ergebnisse auch bei unterschiedlichem Ausbildungsniveau der Untersuchungsgruppe lieferte. Deshalb entfielen auch handelsübliche, gut standardisierte und geeichte Schultests als Untersuchungsmaterial, denn Leistungstests in Rechtschreiben sind grundsätzlich nur als Lückentests aufgebaut - ein Verfahren, das im normalen Rechtschreibunterricht nur sporadisch geübt und in Klassenarbeiten gar nicht benutzt wird (Rechtschreibtest RST 8 von DAMM, HYLLA, SCHÄFER, DRT 4/5 von MEIS, DRT 3/4 von MÜLLER). Rechentests wie der RT 8+(FISCH, HYLLA, SÜLLWOLD) enthalten nur zwei Parallelformen (A und B), d.h. sie waren ungeeignet,

weil drei Varianten von Prüfungssituationen untersucht werden sollten.

GLATZ bietet als Ergebnis der 'Arbeitsgemeinschaft für Leistungsmessung in der Volksschule' zahlreiche Vergleichsarbeiten im Rechnen und Rechtschreiben für das 3. - 9. Schuljahr (Tellus-Verlag) an. Er verspricht in einem Vorwort "Jahre hindurch vorgeprüfte und geeichte Diktate und Rechenarbeiten mit empirisch ermittelten Schwierigkeitsindizes und durchschnittlichen Fehler- und Punktzahlen." Insgesamt arbeiteten 5000 Schüler an 146 Schulen mit. Die Schulen wurden so ausgewählt, daß Repräsentanz der Ergebnisse für das gesamte Bundesgebiet gewährleistet ist. Es wurden knappe Versuchsanweisungen, die identische Instruktionsübermittlung und Vergleichbarkeit der Ergebnisse sichern sollen, erstellt.

Allerdings fehlen nähere Angaben über Größe und Eigenart der Stichprobe, d.h. letztlich der Nachweis der Repräsentanz, wie auch INGENKAMP (1963) bemerkte, obwohl er ihren Wert als Vergleichsarbeiten betonte. Der Verfasser fand keine anderen Vergleichsarbeiten, die sich besser als Versuchsmaterial eigneten.

Der Diktatversuch verlangte vom Probanden die exakte Reproduktion von gespeicherten, früher gelernten Wortbildern. Es handelte sich vor allem um Gedächtnisleistungen. Da aber gleichzeitig Satzteile verstanden und behalten und Satzzeichen gesetzt werden mußten, entschied wohl auch Konzentrationsfähigkeit die Höhe der Fehlerrate. Der Inhalt aller Diktate war ohne Schwierigkeiten zu erfassen.

Das Zahlenrechnen prüfte die eindeutige Beherrschung eingeübter Regeln und Lösungsschemata. Da aber nach jeweils drei Aufgaben neue oder zumindest abgewandelte

Rechenverfahren gefordert wurden, kam es für die Schüler darauf an, sich auf die neuen Verfahren einzustellen und nicht allzu blind und sorglos vorzugehen.

2.5.1. Diktate

Diktat 9. Schuljahr

Die für das 9. Schuljahr vorgesehenen Diktate erwiesen sich entsprechend der notierten Fehlerquote als wenig schwierigkeitsidentisch. Das Versuchsmaterial wurde deshalb dem Angebot des 8. Schuljahres entnommen.

Diktate (Angaben nach GLATZ)

a. Übung: Der Güterverkehr		164 Wörter
Durchschnittliche Fehlerzahl	7,6	
Rechtschreibung	5,2	
Zeichensetzung	2,4	
b. Übung: Der Elektromotor		181 Wörter
Durchschnittliche Fehlerzahl	5,1	
Rechtschreibung	2,9	
Zeichensetzung	2,2	
c. Hauptversuch: Diktat A: Die klügsten Bauern haben die größten Kartoffeln		224 Wörter
Durchschnittliche Fehlerzahl	9,1	
Rechtschreibung	6,4	
Zeichensetzung	2,7	
d. Hauptversuch: Diktat B: Internationaler Fernsprechverkehr		118 Wörter
Durchschnittliche Fehlerzahl	9,0	
Rechtschreibung	6,8	
Zeichensetzung	2,2	

e. Hauptversuch: Diktat C: Das Watt 199 Wörter

Durchschnittliche Fehlerzahl	8,9	
Rechtschreibung	5,6	gemittelt
Zeichensetzung	3,3	Jungen u.Mädchen

Diktate 4. Schuljahr

Auf Grund der Ergebnisse von Vorversuchen wurden aus einem Angebot von sechs Texten drei Erzählungen für den Hauptversuch ausgewählt. Ein Diktat (Der Sämann), das sich auf Grund der formalen Kriterien (Durchschnittsfehlerzahl und Textlänge) an sich eignete, wurde wegen Einsprüche zahlreicher Lehrer, die den Inhalt der Erzählung kritisierten, gestrichen und durch einen anderen Text ersetzt (Räuber und Polizist).

Diktate (Angaben nach GLATZ)

a. Übung: Die hungrige Drossel 95 Wörter
Durchschnittliche Fehlerzahl 4,3

b. Übung: Ein kleiner Freund 103 Wörter
Durchschnittliche Fehlerzahl 4,7

c. Hauptversuch Diktat A: Eine Igelgeschichte 128 Wörter
Durchschnittliche Fehlerzahl 5,6

d. Hauptversuch Diktat B: Räuber und Polizist 110 Wörter
Durchschnittliche Fehlerzahl 6,2

e. Hauptversuch Diktat C: Am Bienenstand 131 Wörter
Durchschnittliche Fehlerzahl 5,7

Wichtigstes Kriterium bei der Auswahl der Vergleichsarbeiten war die durchschnittliche Fehlerzahl als Indiz für die Schwierigkeit der Diktate. So konnten allerdings Unterschiede in der Textlänge nicht vermieden werden. Das zweite Kriterium im 9. Schuljahr war die Übereinstimmung des Fehlerverhältnisses Rechtschreibung - Zeichensetzung.

Leistungsbewertung

Aus Verrechnungsgründen wurden nur ganze Fehlerpunkte angerechnet, wobei vergessene Oberzeichen (i-Punkte, ö-, ä- und Ü-Striche) nicht berücksichtigt waren. Zeichensetzungsfehler wurden als ganze Fehler bewertet.

Enthielt ein Wort mehrere orthographische Verstöße, wurde nur ein Fehler angestrichen. Ein Wort, das wiederholt den gleichen Verstoß aufwies, zählte nur einmal im Verlauf des Diktates als Fehler. Ausgelassene, zuviel oder ersatzweise geschriebene Wörter ergaben je einen Fehler (nach GLATZ).

Erkundungsexperimente (HENSCHEL 1971, NOAK 1972) zeigten, daß eine Aufgliederung der Rechtschreibfehler in mehrere Kategorien (Rechtschreibfehler R, Auslassungsfehler A, Oberzeichenfehler O) keine neuen Erkenntnisse für das vorliegende Problem lieferte. Das Fehlerverhältnis (R - A - O) schien sich nur geringfügig unter dem Einfluß der experimentellen Bedingungen zu ändern. Alle orthographischen Fehler wurden daraufhin als W-Fehler, Zeichensetzungsfehler als Z-Fehler notiert.

2.5.2. Rechnen

Wie sich aus den Vorversuchen ergab, kam es im schriftlichen Zahlenrechnen darauf an, Aufgaben zu finden, deren Lösungswege den Schülern bekannt waren, um in etwa identische Ausgangsbedingungen zu gewährleisten. So wurden Stoffbereiche herangezogen, die entweder mindestens ein oder zwei Monate vor dem Hauptversuch (4.Schuljahr) erarbeitet oder im Laufe von Wiederholungen neu aufbereitet worden waren (9. Schuljahr: Gemeine Brüche, Dezimalbrüche). Problemaufgaben, deren Lösungswege Einblick in

die höheren Prozesse des diskursiven und logischen Denkens, der Umstrukturierung und der Neuorientierung (WERTHEIMER, DUNCKER, METZGER) geben konnten, wurden nicht als Prüfungsmaterial verwendet, da die notwendigen statistischen Kriterien zur Entwicklung von drei Parallelserien nur mit riesigem Aufwand zu bestimmen waren. Außerdem waren Kinder aus vielen Schulen zu untersuchen, in deren Unterrichtsgestaltung die Lösung von Problemaufgaben unterschiedliches Gewicht hatte. Deshalb wurden nur Aufgaben des schriftlichen Zahlenrechnens von mittlerer bis geringer Schwierigkeit ausgewählt, die in erster Linie Rechenfertigkeit prüften.

Aus einem Rechenproblem (z.B. Teilen einer gemischten Zahl durch einen Bruch) wurden nach Absprache mit interessierten und qualifizierten Lehrern neun schwierigkeits- und strukturidentische Aufgaben entwickelt. Nach dem Zufallsprinzip zugeordnet ergab sich für die Versuchsserien A, B, C je eine Aufgabengruppe mit drei Aufgaben.

Für das 4. Schuljahr wurden die Parallelserien aus vier unterschiedlich schwierigen Rechenproblemen entwickelt (= 12 Aufgaben je Serie). Im 9. Schuljahr bildeten fünf Rechenprobleme mit steigender Schwierigkeit (= 15 Aufgaben je Serie) die Grundlage für das Aufgabengerüst.

Zahlenrechnen 9. Schuljahr

Die Vergleichsarbeiten von GLATZ gaben für bestimmte Rechenprobleme empirisch gesicherte Durchschnittspunktzahlen an. Nach Überprüfung dieser Werte (und notwendiger Korrekturen) an rund 200 Schülern wurde ein Aufgabengerüst mit leicht ansteigendem Schwierigkeitsgrad (je Stufe drei Aufgaben) erstellt.

1. Rechenproblem: Dezimalbruch mal Dezimalbruch
   Aufgaben 1 - 3, Beispiel: 19,07 . 70,09 p = 0,87[1)]
2. Rechenproblem: Dezimalbruch geteilt durch Dezimalbruch
   Aufgaben 4 - 6, Beispiel: 0,17334 : 0,117 p = 0,70
3. Rechenproblem: Addieren echter Brüche
   Aufgaben 7 - 9, Beispiel $\frac{3}{4} + \frac{7}{8} + \frac{1}{6}$ p = 0,68
4. Rechenproblem: Gemischte Zahl geteilt durch einen echten Bruch
   Aufgaben 10 - 12, Beispiel: $1\frac{5}{6} : \frac{1}{2}$ p = 0,58
5. Rechenproblem: Gemischte Zahl geteilt durch gemischte Zahl mal Dezimalbruch
   Aufgaben 13 - 15, Beispiel: $4\frac{2}{7} : 1\frac{1}{14} \cdot 0{,}2$ p = 0,47

Zahlenrechnen 4. Schuljahr

Die Rechenaufgaben der Vergleichsarbeiten konnten nur bedingt herangezogen werden, da ein Teil der Aufgabentypen lt. Lehrplangestaltung des Landes NRW erst gegen Ende des Schuljahres behandelt wird. So wurden in Absprache mit Lehrern und mit Hilfe von Voruntersuchungen angemessene Rechenprobleme herausgestellt. Die statistische Kontrolle ergab, wie im 9. Schuljahr, mittlere bis geringe Schwierigkeitsgrade.

1. Rechenproblem: Subtrahieren ganzer Zahlen mit Zehnerübergang
   Aufgaben 1 - 3, Beispiel: 667 - 239 , p = 0,88
2. Rechenproblem: Malnehmen ganzer Zahlen
   Aufgaben 4 - 6, Beispiel: 53 . 49 DM p = 0,62

[1)] Schwierigkeitsgrad der Aufgaben, ermittelt in einer Kontrolluntersuchung mit 200 Schülern.

3. Rechenproblem: Addieren mehrerer ganzer Zahlen
   Aufgaben 7 - 9, Beispiel: 5127 + 40 + 901 + 8 + 4129 p = 0,85
4. Rechenproblem: Addieren und Subtrahieren ganzer Zahlen
   Aufgaben 10 - 12, Beispiel: 5586 - 77 - 631 + 4197 p = 0,49

Leistungsbewertung

Jede falsche Lösung wurde mit einem Fehlerpunkt bewertet. Richtige Lösungen mit falschem Lösungsweg galten als Fehler. Höchste Fehlerzahlen: im 9. Schuljahr 15 Fehler; im 4. Schuljahr: 12 Fehler.

2.5.3. Analyse der Fragebogen

Es sollte eine Frageform gefunden werden, die die Einstellung des einzelnen Schülers zu einem bestimmten Schulfach (hier: Rechtschreiben oder Rechnen) festlegt. Nach GUILFORD (1964) versteht man unter Einstellung die Neigung eines Menschen, bestimmten sozialen Objekten oder Aktionen gewogen oder nicht gewogen zu sein. Als soziale Aktionen können auch schulische Aktionen angesehen werden, wie Prüfungsarbeiten oder Einstellungen zu bestimmten Fächern. Nach GUILFORD schließt eine Einstellung Überzeugungen und Gefühle ein. Indikatoren für Einstellungen sind z.B. Meinungen, die eine Person äußert. THURSTONE versuchte 1928 mit einer quantitativen Methode, Einstellungen zu messen. Er benutzte ein Einstellungskontinuum, das sich von einem positiven Extrem über Gleichgültigkeit zum negativen Extrem erstreckt. Weiterführende Voraussetzungen dieser Methode (vgl. die zusammenfassenden Äußerungen bei GUILFORD 1964) sollen hier nicht diskutiert werden.

Da der Verfasser keine standardisierten und validierten

Einstellungsfragebogen bezüglich seiner Thematik vorfand, mußte er zur Grobdifferenzierung der Schülereinstellungen gegenüber einem bestimmten Schulfach ein Verfahren zu Hilfe nehmen, dessen Stabilität durch dreifache Einstellungsschätzung im Zeitraum der Untersuchung erhöht werden sollte.

a. Indirekte Befragung. Die Schüler schätzen das Ergebnis einer fiktiven experimentellen Untersuchung.

b. Direkte Befragung. Schüler notieren ihre eigene Meinung zum betreffenden Schulfach.

c. Direkte Befragung des Lehrers als Schätzung der Einstellung des Schülers.

Nach wenig ergiebigen Vorversuchen mit einer fünfstufigen Skala wurde eine dreistufige Einteilung verwendet (gerne haben, gerne mögen - + -,/ gleichgültig, ich weiß nicht - 0 -,/ ich mag nicht - - -).

Die Gesamtergebnisse der Einstellungsschätzung konnten nur für den Zeitraum der Untersuchung von Bedeutung sein, da Einstellungen nicht losgelöst von den sich immer wieder ändernden Bedingungen im pädagogischen Feld betrachtet werden können.

Alle Gesamtergebnisse, die sowohl extrem positive als auch negative Schätzungen aufwiesen (z.B. + - +, - + -), deren Schätzungen überwiegend der Mitte zwischen den Polen einzuordnen waren (z.B. 0 - 0, 0 0 0), und die sich weniger als auf drei Schätzungen stützten, wurden ausgeschieden. Nur solche Schüler, deren Unterlagen eindeutig negative oder positive Einschätzungen auswiesen, fanden Berücksichtigung.

Dieses Verfahren kann infolge fehlender empirischer Überprüfung höchstens als grobe Schätzung der Einstellung betrachtet werden.

2.6. Anordnung und Durchführung der Versuche

Alle Aufgaben wurden in der Form des Gruppenversuchs gegeben. Der Hauptversuch nahm einschließlich der Instruktionen 3 mal 45 Minuten Zeit in Anspruch. Die Einführungs- und Übungsphase erforderten unter Einbeziehung aller Nebenaufgaben noch einmal 4 mal 45 Minuten Zeit. In Vorversuchen hatte es sich gezeigt, daß es unbedingt notwendig war, eine vertrauensvolle, relativ entspannte, leistungsdruckarme Arbeitsatmosphäre für den Hauptversuch (Übungsarbeit) herzustellen. Dies erforderte eine Eingewöhnungsperiode von zwei Unterrichtsstunden, in denen der Versuchsleiter als Lehrer den geplanten Unterricht im betreffenden Fach weiterführte. Danach konnte gewöhnlich (bis auf zwei Ausnahmen) in zwei weiteren Stunden die Aufbereitung des Prüfungsmaterials im Zahlenrechnen (Abbau von Übungseffekten) erfolgen oder die Vorbereitung der Prüfdiktate (Gewöhnung an die vorher festgelegte Diktiermethode bzw. an die Eigenheiten des Versuchsleiters während des Diktates) vorangetrieben werden.

Folgende Voraussetzungen mußten vor Beginn des Hauptversuches erfüllt sein:

- Anerkennung des Versuchsleiters als vollwertiger Lehrer,
- Konfliktlose Beziehung zwischen Versuchsleiter und den Schülern,
- Einbettung des Versuchs in den üblichen Schulunterricht,
- Unwissenheit der Schüler hinsichtlich Methode und Ziel der Untersuchung,
- Ausführen von Übungsarbeiten ohne Verdacht der nachträglichen Lehrerkontrolle,

- Vermeiden jeglichen Zeitdrucks.

Wenn diese Voraussetzungen nicht erfüllt waren, wurde der Hauptversuch in der betreffenden Klasse nicht begonnen und der Gesamtversuch abgebrochen, da die Identität der Bedingungen der Ausgangslage nicht gewährleistet war.

Außerdem wurde streng darauf geachtet, daß
- wenigstens zwei, höchstens aber vier Tage zwischen den einzelnen Versuchen lagen,
- schulische Ereignisse wie Schulfeiern, Sportwettkämpfe, Ausstellungsbesuche den Ablauf des Gesamtversuchs nicht beeinträchtigten,
- die Untersuchung nur in der Zeit von ca. 8.45 Uhr - ca. 11.45 Uhr durchgeführt wurde (also nicht in der ersten, fünften oder sechsten Stunde, montags und samstags nur in Ausnahmefällen).

Im Hauptversuch wurde die Reihenfolge der Diktate A,B,C und der Rechenarbeiten A, B, C von Klasse zu Klasse so geändert, daß jede Arbeit gleich oft an erster, zweiter und dritter Stelle in der Versuchsanordnung stand (A, B, C; B, C, A; C, A, B).

Übungssituation

Nun wurden die Schüler gebeten - wie schon in der Eingewöhnungszeit -, ihre Übungshefte zu öffnen und die vorgedruckten Aufgaben zu lösen (Zahlenrechnen) oder das angebotene Diktat unter den schon gewohnten Bedingungen niederzuschreiben. Der Lehrer wies die Schüler darauf hin, daß man selbständig arbeiten könne, um zu erkennen, ob man den Anforderungen gewachsen sei oder nicht. Schüler, die sich nicht an diesen

Hinweis hielten, wurden mit freundlichen, verständnisvollen Worten noch einmal eingewiesen. Die Übungsarbeiten sollten in möglichst unbelasteter, leistungsdruckarmer Atmosphäre ausgeführt werden. Der geringste Hinweis auf nachfolgende Leistungskontrolle durch den Lehrer unterblieb. Alle üblichen Verhaltensweisen des Lehrers während der Durchführung von Klassenarbeiten wie Drohen, Verbieten, Tadeln, Entzug der Zuneigung wurden vermieden.

Nachdem alle Schüler ihre Übungsarbeit beendet hatten oder zumindest bei weiteren Versuchen keine nennenswerten Verbesserungen zu erzielen glaubten, wurden die Schüler aufgefordert, ihren Namen auf die Übungsblätter zu schreiben und sie aus dem Heft herauszutrennen. Der Versuchsleiter sammelte mit einigen Schülern als Helfer die Blätter schnell ein, damit nicht nachträglich Veränderungen vorgenommen werden konnten. Die zumeist etwas aufgeregten Schüler wurden mit dem Hinweis beruhigt, daß der Lehrer nach einer langen Übungsserie wissen muß, ob die Schüler in ihrer Arbeit Erfolg gehabt haben. Hinweise auf Zensuren unterblieben, um das entspannte Lehrer - Schüler - Verhältnis nicht zu belasten.

Klassenarbeitssituation

Die Schüler schrieben diese Arbeit unter gewohnten Klassenarbeitsbedingungen in die dafür vorgesehenen Hefte mit fortlaufender Numerierung oder auf DIN-A-4-Blätter, die in einer Arbeitsmappe gesammelt wurden. Während des Versuchsablaufs war darauf zu achten, daß die gewohnten Lehrerverhaltensweisen die übliche, gespannte Klassenarbeitsatmosphäre hervorriefen. Die Schüler wurden auf die Wichtigkeit der Arbeit hingewiesen. Wie üblich verbot ihnen der Lehrer, beim Nachbarn abzuschauen oder selbst Hilfestellung zu leisten. Während der Klassenarbeit hatte der Versuchs-

leiter darauf achtzugeben, daß die Schüler selbständig arbeiteten und die ausgesprochenen Verbote beachteten.

Nach Beendigung der Arbeit wurden die Hefte eingesammelt.

Prüfungssituation

Die im üblichen Schulunterricht nicht vorgesehene, einmalige Prüfungsarbeit wurde auf DIN-A-4-Blätter niedergeschrieben, die der Versuchsleiter zur Verfügung stellte. Als Überschrift notierten die Schüler 'Prüfungsarbeit'. Die Anweisung war so gehalten, daß die Prüfungsarbeit als in höchstem Maße mitentscheidendes Kriterium für die Benotung in den für den weiteren Lebensweg wichtigen Abschluß- und Abgangszeugnissen ausgewiesen wurde.

Der Lehrer achtete sehr aufmerksam auf die Einhaltung aller ausgesprochenen Verbote. Verstöße wurden ohne vorherige Warnung geahndet.

Im Anschluß an die Prüfungsarbeit gaben die Schüler ihre Einstellung zu dem betreffenden Schulfach wieder. Damit war der Unterrichtsversuch abgeschlossen.

Durchführung der weiteren Datenerhebung

Der zuständige Klassen- oder Fachlehrer gab seinen Eindruck hinsichtlich der positiven oder negativen Einstellung der Schüler zum entsprechenden Fach auf dem angegebenen Kontinuum wieder. Ferner wurden die letzten Zeugniszensuren im betreffenden Fach notiert.

Ein Rechentest (RT 8+) bzw. ein Rechtschreibtest (RST 8+) zur differenzierten Feststellung der Leistungsstärke rundeten im 9. Schuljahr die Datenerhebung ab. Auf Grund der Zeugniszensur und des Testergebnisses wurde

die Rangordnung einer Klasse bezüglich der Leistungsstärke im betreffenden Fach erstellt.

Im 4. Schuljahr konnte ein Schulleistungstest nicht durchgeführt werden, da für das Fach Rechnen kein geeigneter Test (nur ungeeicht) zur Verfügung stand. Deshalb wurden im 4. Schuljahr die Rangplätze bezüglich der Leistungsstärke auf Grund der Zeugniszensur und der Fehlerzahl im Hauptversuch vergeben.

Nach Abschluß aller Versuche wurden die Schüler über Sinn und Zweck des Experiments aufgeklärt und erhielten den für sie zumeist beruhigenden Hinweis, daß das Ergebnis der 'Prüfungsarbeit' keinerlei negative Auswirkungen auf die Zeugniszensuren habe.

Kontrolluntersuchungen im Zahlenrechnen

Der experimentelle Aufbau des Rechenversuchs legte den Schluß nahe, daß Mittelwertsunterschiede zwischen den Meßreihen nicht nur auf die Wirkung der experimentellen Bedingungen zurückzuführen waren, sondern auch infolge Übung, Übersättigung oder ähnlicher zeit- oder wiederholungsabhängiger Faktoren zustande kamen. Die Eigenart der struktur- und schwierigkeitsidentischen Prüfarbeiten erforderte den Vergleich mit Kontrollergebnissen, um den Effekt der intendierten experimentellen Bedingungen von denen der Wiederholung zu trennen. Die Kontrollgruppen wurden in drei Abschnitten unter möglichst identischen Bedingungen mit dem im Hauptversuch verwendeten Aufgabenmaterial untersucht. Im Zeitraum zwischen den Untersuchungen durften keine stimulierenden Auskünfte negativer oder positiver Art gegeben werden. Im übrigen galten die Bedingungen und Voraussetzungen des Gesamtversuchs.

2.7. Anweisungen (Beispiel)

Zahlenrechnen 4. Schuljahr

Erste Phase der Vorbereitung des Hauptversuchs

"Ich habe euch einige vorgedruckte Übungsaufgaben mitgebracht, die ihr wahrscheinlich ohne große Schwierigkeiten lösen könnt. Nehmt bitte eure Übungshefte (oder Vorschreibhefte) heraus. Da die Aufgaben leicht sind, braucht ihr erst gar nicht abzuschauen. Außerdem wollt ihr doch sicherlich wissen, ob ihr die Aufgaben lösen könnt.

Ihr könnt die Aufgaben so rechnen, wie ihr es gelernt habt. Ich werde die Hefte nach getaner Arbeit nicht einsammeln, sondern ihr werdet die Hefte mit dem Nachbarn austauschen. Wir werden dann zusammen die Aufgaben kurz durchgehen. Ihr notiert bitte die richtigen Lösungen des Nebenmannes. Versucht, die Aufgaben möglichst richtig zu lösen. Die Zeit spielt für die Bewertung eurer Leistung keine Rolle. Es kommt also nicht auf die Zeit an, sondern allein auf die selbständige Arbeit."

Nach Beendigung der Arbeit (ca. 20 - 30 Minuten):

"So, nun werden wir die Aufgaben der Reihe nach durchgehen. Wer noch Schwierigkeiten in der Ausrechnung hat, möge es bitte mitteilen."

Zum Schluß:

"Wie ihr gesehen habt, haben manche von euch noch Lücken, obwohl andere fast fehlerfrei gearbeitet haben. Wir werden auch in den nächsten Stunden noch einige Aufgaben rechnen."

Zweite Phase der Vorbereitung des Hauptversuchs

"Ich habe euch wieder einige Aufgaben auf diesen Blättern mitgebracht. Sie sind denen, die wir in der letzten Stunde gerechnet haben, ähnlich. Ihr wißt ja, wie wir das machen. Zeit habt ihr genug. Später kontrolliert der Nachbar wieder die Ergebnisse. Die Aufgaben sind nicht schwer. Jeder kann sie lösen, ohne beim Nachbarn abzuschauen."

Übungsarbeit

"Die heutigen Übungsaufgaben rechnet ihr wieder wie gewohnt." (Sonstige Hinweise fehlen. Partnerschaftskontrolle wird nicht versprochen. Vorversuche ergaben, daß Rückfragen in dieser Richtung nicht zu erwarten sind.)

Nach Abschluß der Aufgabenserie:

"So, nun möchte ich mal sehen, was ihr nach dieser Übungsserie leisten könnt. Trennt bitte die Seiten mit den gerechneten Aufgaben aus eurem Heft heraus und schreibt euren Namen und Vornamen darauf. Nun sammelt in der gewohnten Weise schnell ein."

Klassenarbeit

"Wir werden heute eine Klassenarbeit schreiben. Es werden wieder Aufgaben gegeben, die euch nicht ganz unbekannt sind. Zeit habt ihr genug. Die Arbeit ist ziemlich wichtig für euch. Deshalb kann ich kein Abschauen gestatten." (keine weiteren Hinweise)

Prüfungsarbeit

" Wir werden heute eine einmalige Prüfungsarbeit, die für euch äußerst wichtig ist, schreiben. Wie ihr wißt, werdet ihr die Grundschule am Ende dieses Schuljahres verlassen, um entweder das Gymnasium, die Realschule oder die Hauptschule zu besuchen. Eure Lehrer müssen ganz genau wissen, was ihr könnt, deshalb auch diese wichtige Prüfungsarbeit. In den vergangenen Tagen und Wochen hattet ihr Zeit und die Möglichkeit, euch mit den Aufgaben, die ihr in dieser Prüfungsarbeit wiederfindet, zu befreunden.

Strengt euch mächtig an, reißt euch zusammen, damit ihr ein gutes Ergebnis erzielt.

Vorher noch eine Warnung: Demjenigen, der beim Abschauen erwischt wird, wird seine Arbeit sofort abgenommen und mit 'sechs' zensiert.

Bevor ihr anfangt, schreibt als Überschrift 'Prüfungsarbeit', wie es an der Tafel steht, und oben links in die Ecke des Blattes euren Vornamen, den Hausnamen und den Namen der Schule."

Die umfangreichen Versuchsanweisungen des 4. Schuljahres Rechtschreiben, des 9. Schuljahres und der weiteren Datenerhebung sind wegen ihrer Ausführlichkeit im Anhang nachzulesen.

3. EINFLUß DER PRÜFUNGSBEDINGUNGEN

Vorbemerkung zur Datenverarbeitung

Die Schülerzahl mußte reduziert werden, weil diejenigen Schüler, von denen wichtige Daten nicht zu erhalten waren, ausgeschlossen wurden, ohne daß sich soziologisch und pädagogisch entscheidende Stichprobenverhältnisse grundlegend änderten.

Vorbemerkung zur Auswahl der statistischen Prüfverfahren

Die Leistungen der Probanden wiesen in ihrem Streuungsbereich meist keine Normalverteilung auf (besonders in den Leistungsabschnitten des Gesamtbildes im Zahlenrechnen); deshalb mußten parameterfreie Prüfverfahren zum Vergleich der Leistungen in den verschiedenen Prüfungssituationen herangezogen werden. Zunächst wurde die von FRIEDMAN (1937) entwickelte Rangvarianzanalyse[1)] benutzt, um Einflüsse der Prüfungssituationen in dem nach Art sogenannter KENDALL-Pläne aufgebauten Experiment festzustellen. Ein signifikanter Chi-Quadrat-Wert im FRIEDMAN-Test sagt nichts über die Richtung von Leistungsdifferenzen aus und gibt auch nicht an, welche Unterschiede zwischen mehreren korrelierenden Stichproben statistisch gesichert werden können, sondern prüft nur, ob die experimentellen Bedingungen einen signifikanten Einfluß auf die Verteilung der entsprechenden Meßwerte ausüben. Daher muß zur eingehenderen Analyse ein Verfahren benutzt werden, das den Vergleich zweier korrelierender Stichproben von Meßwerten gestattet. Ein geeignetes Verfahren ist der WILCOXON-Test für Paardifferenzen, eine

[1)] Berechnung nach E. WEBER, Grundriß der biologischen Statistik, 1967 und G.A. LIENERT, Verteilungsfreie Methoden in der Biostatistik, 1962

Kombination von Rangtest und Vorzeichentest, den LIENERT (1962) als effizientes Prüfmittel bezeichnet (asymptotische Effizienz von E=0,95).

Infolge großer Stichprobenumfänge und des dadurch bedingten Rechenaufwandes mußte die Anlage des Rechenzentrums der Universität zu Karlsruhe (Univac 1108) in Anspruch genommen werden.[1] Weniger zeitraubende Datenanalysen wurden per Tischrechner ausgeführt.

Vorbemerkungen zur Befunddarstellung

Es war nicht nur eine Analyse des Gesamtbildes der Leistungsausprägung vorgesehen, sondern auch eine differenzierte Untersuchung des Leistungsverlaufs. Im folgenden wird zunächst das Gesamtergebnis untersucht, danach der Leistungsablauf. Während im Zahlenrechnen nur ein Fehlerwert zu berücksichtigen ist (pro falsche Lösung 1 Fehlerpunkt), kommen in den Diktatversuchen zwei Fehlerarten zur Anrechnung: einmal die W-Fehler, die das Fehlen oder die Falschschreibung eines Wortes anzeigen, zum anderen die Z-Fehler als Zeichensetzungsfehler. In der folgenden Befunddarstellung sind zunächst nur Gesamtfehlerhäufigkeiten (W+Z) notiert; auf die Differenzierung nach W- und Z-Fehler wird in einem späteren Kapitel eingegangen.

3.1. Leistungsunterschiede im Gesamtergebnis beider Schuljahre

Als erstes werden die Gesamtergebnisse miteinander verglichen, um globale Tendenzen hinsichtlich der Einwirkung von Prüfungssituationen auf das Leistungsverhalten von Schülern im 4. und 9. Schuljahr konstatieren zu können.

[1] Ich danke Herrn W. Schmitz für wertvolle Hilfen beim Programmieren der anstehenden Probleme.

Tabelle 1: Einfluß von Prüfungssituationen im Gesamtergebnis beider Schuljahre (Prüfverfahren: FRIEDMAN-Test)

Schuljahr	Fach	Vpn	Chi-Quadrat-Wert
4.	Rechnen	358	3.204
			>.10*
4.	Diktat	234	2.932
			>.10
9.	Rechnen	324	8.528
			<.05
9.	Diktat	201	37.545
			<.01

* Zufallswahrscheinlichkeit der Chi-Quadrat-Werte im Friedman-Test

Eine Prüfung der Rangsummenunterschiede mit dem FRIEDMAN-Test läßt erkennen, daß im 4. Schuljahr die experimentellen Bedingungen keinen wesentlichen Einfluß auf die Verteilung der Meßwerte haben. Die Leistungsverteilungen im 9. Schuljahr dagegen sind so eindeutig, daß die Wirkung systematischer Einflüsse von Prüfungsbedingungen angenommen werden muß. Bemerkenswert ist, daß die von den Schülern empfundene unterschiedliche Schwierigkeit des Prüfmaterials die Höhe der Chi-Quadrat-Werte nicht in bedeutsamem Maße ändert. Die Rechenaufgaben werden nicht zuletzt auf Grund der Stoffaufbereitung vor dem Hauptversuch als recht gut lösbar und als eigentlich im ganzen nicht schwierig erlebt (sowohl die Lösungsfrequenzen der Aufgaben in der Voruntersuchung als auch die geringe Anzahl der Arbeiten mit hohen Fehlerquoten - 13,14 oder 15 Fehlern - untermauern diesen Tatbestand), während die durchschnittliche Fehlerzahl in den Diktatversuchen wesentlich höher liegt als vergleichbare Durchschnittsfehleran-

gaben aus den letzten Klassenarbeitsdiktaten. Trotzdem weist der Diktatversuch im 4. Schuljahr ebensowenig bedeutsame Leistungsunterschiede auf wie die als leicht empfundenen Rechenaufgaben (4.Schuljahr), dagegen ist im 9. Schuljahr sowohl bei leichten Rechenaufgaben als auch bei schwierigen Diktattexten Belastungseinfluß mehr (1%-Niveau) oder weniger (5%-Niveau) deutlich nachweisbar.

Es kann also festgehalten werden, daß nicht so sehr das Fach oder die Schwierigkeit der gestellten Aufgabe für statistisch gesicherte Rangsummenunterschiede der Meßwertverteilungen verantwortlich ist, sondern in erster Linie das unterschiedliche Alter und damit eng verknüpft die Unterrichtserfahrung.

Da der FRIEDMAN-Test hinsichtlich der Analyse detaillierter Sachverhalte nur wenige Ergebnisse liefert, soll nun das Verhältnis einzelner Meßreihen mit dem WILCOXON-Test eingehender untersucht werden. Das WILCOXON-Prüfungsverfahren verlangt die Bildung von Paardifferenzen. Es werden folgende Leistungsunterschiede geprüft:

Übungsarbeit - Klassenarbeit (ÜA - KA),
Klassenarbeit - Prüfungsarbeit (KA - PA),
Übungsarbeit - Prüfungsarbeit (ÜA - PA).

Tabelle 2 gibt die einzelnen Leistungsdifferenzen und die Zufallswahrscheinlichkeiten der U-Werte der WILCOXON-Prüfung an.

Tabelle 2: Einfluß der Prüfungssituationen auf die Leistungsdifferenzen zwischen ÜA - KA, KA - PA und ÜA - PA (Prüfverfahren: WILCOXON-Test)

Schuljahr	Fach	Vpn	Vergleich ÜA - KA***	Vergleich KA - PA	Vergleich ÜA - PA
4.	Rechnen	358	+0,09**	-0,51	-0,42
			>.10*	<.01	>.05
4.	Diktat	234	+0,05	-0,77	-0,72
			>.05	<.05	<.05
9.	Rechnen	324	+0,33	-1,27	-0,94
			>.05	<.01	<.05
9.	Diktat	201	+3,48	-2,96	+0,52
			<.01	<.01	>.10

* Zufallswahrscheinlichkeit der U-Werte im Wilcoxon-Test

** Die durchschnittlichen Leistungsunterschiede weisen Vorzeichen auf. (-) Fehlererhöhung in der zweitgenannten Arbeit, (+) Fehlerverbesserung in der zweitgenannten Arbeit.

*** ÜA bedeutet im weiteren Verlauf Übungsarbeit, KA bedeutet im weiteren Verlauf Klassenarbeit, PA bedeutet im weiteren Verlauf Prüfungsarbeit.

Vergleicht man die ersten beiden Tabellen, so überrascht, daß auch dort, wo die Chi-Quadrat-Werte des FRIEDMAN-Testes keinen systematischen Einfluß von Prüfungsbedingungen nachweisen, die U-Werte des WILCOXON-Testes Leistungsdifferenzen als statistisch gesichert angeben. LIENERT (1962) verweist auf die unterschiedliche Effizienz der Tests. Die Ergebnisse beider Prüfverfahren liefern insgesamt zuverlässige Angaben zur Sicherung der beobachteten Zusammenhänge.

Nach Tabelle 2 treten die Unterschiede zwischen KA und PA bedeutend stärker und eindeutiger in der Tendenz hervor als die Differenzen ÜA - KA und ÜA - PA. Durchgängig

verschlechtern sich alle untersuchten Schüler in der ungewohnten Prüfungsarbeitssituation (PA), wenn man deren Ergebnisse mit den entsprechenden Klassenarbeitsdaten vergleicht. Die Leistungsminderung im Diktatversuch des 4. Schuljahres kann als leidlich gesichert gelten (5%-Niveau), während die übrigen Leistungseinbußen sehr deutlich (1%-Niveau) ausfallen.

Auch an dieser Stelle wird darauf verwiesen, daß keine entsprechenden Daten gefunden werden können, die einen Einfluß der Schwierigkeit des Prüfmaterials bestätigen. Sowohl als leicht als auch als schwer empfundene Aufgaben bewirken im Vergleich KA - PA durchgängig beobachtete Leistungseinbußen. In erster Linie löst also wohl die Verstärkung der psychischen Belastung durch Instruktion eine Erhöhung der Fehlerzahlen im Vergleich KA - PA aus. Die durch Instruktion als einmalig und bedeutungsvoll angekündigte Prüfungsarbeit verschlechtert das Fehlerniveau vor allem gegenüber der gewohnten studienbegleitenden Klassenarbeit.

Die Tabelle 2 informiert weiterhin über die Leistungsunterschiede ÜA - KA und ÜA - PA. Aus dem Vergleich ÜA - KA können nicht so eindeutige Ergebnisse gefolgert werden. Schüler des 9. Schuljahres schneiden in Übungsdiktaten zwar deutlich schlechter ab als in Klassenarbeiten; die Schüler der anderen Gruppen dagegen zeigen in etwa identische Leistungen in Übungs- und Klassenarbeiten. Dementsprechende nicht signifikante Prüfwerte ($p > .10$) weisen diesen Sachverhalt aus. Die Anspannung während der Klassenarbeit führt also nur in der spezifischen Arbeitssituation des Diktatversuchs der älteren Schüler (9. Schuljahr) zu eindeutigen Leistungsgewinnen

gegenüber Arbeitsergebnissen in der Übungssituation. Im 4. Schuljahr bringt der Druck der Klassenarbeit keine nennenswerten Verbesserungen im Endergebnis.

Da die beobachteten Leistungsunterschiede im Vergleich ÜA - PA hinsichtlich Richtung und Eindeutigkeit uneinheitlich erscheinen, können diese Ergebnisse nur mit Vorsicht zur Befundanalyse herangezogen werden. Im Diktatversuch des 4. Schuljahres und im Rechenversuch des 9. Schuljahres liegen gesicherte Leistungseinbußen (5%-Niveau) vor, während die übrigen beiden Differenzen als nicht unterschiedlich identifiziert werden können ($p > .05$). Die Prüfungsarbeit (PA) führt in bestimmten Fällen Leistungseinbußen herbei, wenn man Ergebnisse der entspannten Übungssituation als Vergleich heranzieht. Leistungsgewinne der Prüfungsarbeit (PA) aber sind im Vergleich ÜA - PA nicht zu beobachten.

Die graphische Darstellung Abbildung 1 gibt anhand durchschnittlicher Fehlerquoten im Diktat und Rechnen recht anschaulich Auskunft über die oben beschriebenen Sachverhalte.

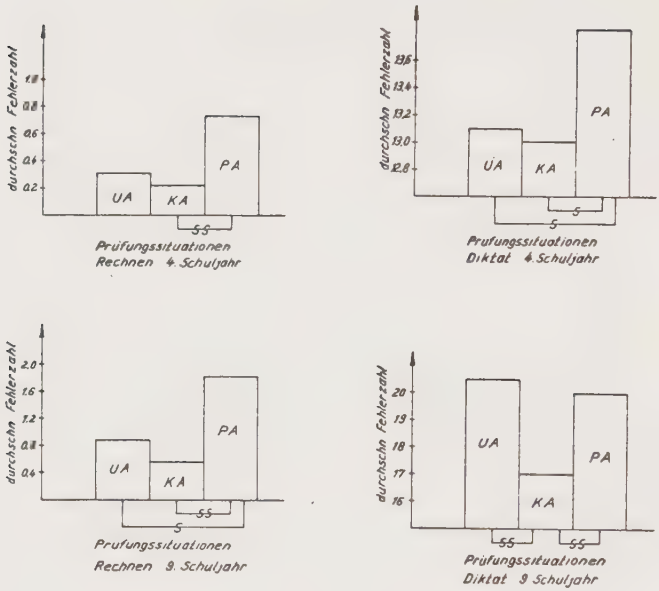

Abb. 1: Graphische Darstellung der Leistungsunterschiede in den drei Prüfungssituationen
Grundlage sind die durchschnittlichen Fehlerzahlen im Diktat und Rechnen. Die Fehlerzahlen im Rechnen wiesen in Einzelfällen negative Vorzeichen auf, nachdem die Fehlerquoten der Kontrolluntersuchung von den Daten der Hauptuntersuchung abgezogen worden waren (lt. Hinweis von WINER 1962). Durch Addition einer Konstanten veränderten sich die Meßreihen so, daß keine negativen Fehlerzahlen beobachtet wurden (Fehlerwerte s. Tab. 3).

Überprüft man die Meßwertverteilung der durchschnittlichen Fehlerquoten aller Schülergruppen (Tab. 3) mit dem FRIEDMAN-Test hinsichtlich ihrer zentralen Tendenz, so rundet der Befund, der eine Wirkung der Prüfungsbedingungen auf die Höhe der Fehlerquoten als in etwa gesichert erscheinen läßt (Chi-Quadrat-Wert: 6,5; $p < .05$), die bisher referierten Ergebnisse ab.

Tabelle 3: Durchschnittliche Fehlerquoten im Gesamtergebnis aller Schülergruppen

Schuljahr	Fach	Vpn	ÜA	KA	PA
4.	Rechnen	358	0,31	0,22	0,73
4.	Diktat	234	13,11	13,06	13,83
9.	Rechnen	324	0,89	0,56	1,83
9.	Diktat	201	20,48	17,00	19,96

Zusammenfassend gewinnt man aus den Leistungsvergleichen der Gesamtergebnisse den Eindruck, daß die experimentellen Bedingungen Veränderungen im Arbeitsergebnis der Schüler hervorrufen. Schüler des 9. Schuljahres werden von diesen Veränderungen stärker betroffen als Viertkläßler. Einfüsse der Facheigenart und Aufgabenschwierigkeit lassen sich nicht nachweisen. Die experimentellen Bedingungen wirken sich am eindeutigsten im Fehlerverhältnis KA - PA aus, und zwar zuungunsten der Prüfungsarbeit. Mit einer Ausnahme (9. Schuljahr, Diktat) schneiden alle Schüler in der relativ belastungsfreien Übungssituation entweder besser ab als in den Belastungssituationen (KA, PA) oder halten das einmal erreichte Fehlerniveau (4. Schuljahr, Rechnen). Die Anspannung in der gewohnten schulischen Leistungsprobe der Klassenarbeit (KA) wirkt sich entgegen mancher Lehrermeinung zumindest im 4. Schuljahr nicht leistungsfördernd aus, wenn man Arbeitsergebnisse der entspannten Übungssituation

als Vergleich heranzieht. Im 9. Schuljahr läßt sich ein Leistungsgewinn in Klassenarbeiten gegenüber Übungsergebnissen nur im Diktatversuch eindeutig nachweisen. Gesteigerte, ungewohnte Belastungen durch Ansprechen wichtiger Ziele wie Versetzung und Erreichen von Berufspositionen seitens eines 'strengen' Lehrers wirken sich eindeutig leistungsmindernd zumindest Klassenarbeitsergebnissen gegenüber aus. Die Schüler schneiden oft auch in Übungsarbeiten (ÜA) besser ab als in den Prüfungsarbeiten (PA). Ein Leistungsvorteil der 'strengen, harten' Instruktionen (PA) kann also letztlich nicht registriert werden.

3.2. Unterschiede im Leistungsverlauf

Neben der Analyse des Gesamtergebnisses ist die des Leistungsverlaufs von besonderem Interesse. Das gesamte Untersuchungsmaterial wurde in Einzelabschnitte aufgespalten und getrennt ausgewertet. Eine Aufteilung in Fünftel (aus Symmetriegründen im Rechenversuch des 4. Schuljahres: Viertel) erhielt den Vorzug vor kleineren Einheiten, die sich in Vorversuchen als wenig aussagekräftig erwiesen hatten, da sie das Verlaufsbild in unübersichtlich kleine Abschnitte aufsplittern (vgl. SANDVOSS 1952, Aufspaltung in Einzelaufgaben). Es hat sich als vorteilhaft herausgestellt (vgl. auch SANDVOSS 1952), große Bereiche des Leistungsverlaufs zusammenzufassen, um eindeutige Aussagen zum Problem der Anfangs- und Endleistung begründen zu können. Deshalb wurden die ersten beiden und letzten beiden Einzelabschnitte zusammengefaßt und als Verlaufsbereich I (Anfangsleistung) und Verlaufsbereich II (Endleistung) analysiert.

Zunächst gilt es aber, die Fehlerverteilung in den Einzelabschnitten (Rechnen: pro Einzelabschnitt drei schwierigkeitsidentische Aufgaben; Diktat: ein Fünftel der Wörter des betreffenden Textes) auf Auswirkung der experimentellen Bedingungen zu überprüfen.

Tabelle 4: Einfluß der Prüfungsbedingungen in den Einzelabschnitten (Prüfverfahren: FRIEDMAN-Test)

Schuljahr	Fach	I. Absch.	II. Absch.	III. Absch.	IV. Abschn.	V.Abschn.
4.	Rechnen	3.573**	7.561	4.031	2.170	--
		>.10*	<.05	>.10	>.10	
4.	Diktat	1.297	1.759	0.660	4.432	6.872
		>.50	>.30	>.50	>.10	<.05
9.	Rechnen	1.985	3.373	6.694	5.596	14.003
		>.10	>.10	<.05	>.05	<.01
9.	Diktat	12.540	15.749	5.157	19.137	11.823
		<.01	<.01	>.05	<.01	<.01

* Zufallswahrscheinlichkeit der Chi-Quadrat-Werte

** Größe der Chi-Quadrat-Werte

Anhand der vorliegenden Daten ist der systematische Einfluß der experimentellen Bedingungen im Leistungsverlauf ersichtlich. Er läßt sich in allen Leistungsverläufen nachweisen, wenn er auch im 4. Schuljahr nur jeweils einen Einzelabschnitt beeinträchtigt und zudem nur auf dem 5%-Niveau gesichert ist. Im Leistungsverlauf des 9. Schuljahres wiegt der Einfluß der experimentellen Bedingungen schwerer. Das Bild des Leistungsablaufes stimmt mit den schon registrierten Tendenzen des Gesamtergebnisses insofern überein, daß gering gesicherte Einflüsse der Prüfungsbedingungen im 4. Schuljahr sehr großen Rangsummenunterschieden im Diktatversuch des 9. Schuljahres gegenüberstehen, die eine

hohe statistische Sicherung der Bedingungseinflüsse beschreiben. Auch die etwas geringere Wirkung der Prüfungsbedingungen auf den Leistungsverlauf in den Rechenarbeiten des 9. Schuljahres wird durch die Größe des Chi-Quadrat-Wertes im Gesamtergebnis ($p < .05$) bestätigt.

Der experimentelle Einfluß scheint sich in den letzten Einzelabschnitten am stärksten auszuwirken, denn drei von vier Meßwertvergleichen weisen den systematischen Einfluß der situativen Bedingungen aus, zwei davon in erhöhtem Maße (1%-Niveau). Zwar lassen sich in jedem Einzelabschnitt experimentelle Einflüsse nachweisen, doch schwerpunktartige Einwirkungen können mit Ausnahme des Schlußabschnittes nicht ausgemacht werden.

Zusätzliche Hinweise bezüglich quantitativer Veränderungen unter dem Einfluß der einzelnen Probesituationen liefert eine Prüfung der Leistungsunterschiede mit dem WILCOXON-Test.

Tabelle 5: Unterschiede im Leistungsverlauf
(Prüfverfahren: WILCOXON-Test)

Schuljahr	Fach	Vergleich	Abschnitt I	II	III	IV	V
4.	Rechnen	ÜA - KA	-0,11	-0,17	+0,24	+0,12	--
			>.05*	>.05	<.05	>.05	
4.	Rechnen	KA - PA	-0,18	-0,24	-0,18	+0,08	--
			<.05	<.05	<.05	>.05	--
4.	Rechnen	ÜA - PA	-0,29	-0,41	+0,06	-0,20	--
			<.01	<.01	>.05	>.05	--
4.	Diktat	ÜA - KA	-0,04	+0,21	-0,14	+0,32	-0,31
			>.10	>.05	>.10	>.05	<.05
4.	Diktat	KA - PA	-0,14	-0,32	-0,06	-0,20	-0,05
			>.10	<.05	>.10	<.05	>.10
4.	Diktat	ÜA - PA	-0,18	-0,09	-0,20	+0,12	-0,36
			>.05	>.10	>.10	>.10	<.05
9.	Rechnen	ÜA - KA	+0,24	-0,18	+0,24	-0,04	+0,07
			<.05	>.10	<.05	>.10	>.10
9.	Rechnen	KA - PA	-0,20	+0,35	-0,46	-0,37	-0,60
			<.05	<.01	<.01	<.01	<.01
9.	Rechnen	ÜA - PA	+0,04	+0,17	-0,22	-0,41	-0,53
			>.10	>.10	<.05	<.01	<.01
9.	Diktat	ÜA - KA	+0,43	-1,14	+0,28	+0,87	+0,78
			<.01	<.01	>.05	<.01	<.01
9.	Diktat	KA - PA	-0,87	-0,60	-0,37	-0,59	-0,53
			<.01	<.01	<.05	<.01	<.01
9.	Diktat	ÜA - PA	-0,44	+0,54	-0,11	-0,28	+0,25
			<.05	<.05	>.10	>.10	>.10

* Zufallswahrscheinlichkeit der U-Werte im Wilcoxon-Test

Man glaubt auf den ersten Blick zu erkennen, daß der Leistungsverlauf von Entlaßschülern der Hauptschule dem Einfluß der experimentellen Bedingungen stärker unterliegt als der von Viertkläßlern, wie auch nach der Prüfung mit dem FRIEDMAN-Test vermutet. Im 4. Schuljahr

können von 27 Leistungsdifferenzen nur zehn statistisch gesichert werden, davon zwei auf dem 1%-Niveau. Die unterschiedlich belastenden Prüfungssituationen beeinflussen den Leistungsablauf im 9. Schuljahr derartig einschneidend, daß 21 von 30 Leistungsunterschieden statistisch bedeutsam sind, davon 14 in erhöhtem Maße (1%-Niveau). Schlüsselt man die Daten auch nach Fächern auf, wird deutlich, daß sich die Schuljahresgruppen sowohl im Rechnen (etwas weniger ausgeprägt) als auch im Diktatversuch hinsichtlich des Anteils an signifikanten Leistungsdifferenzen unterscheiden ($p < .05$). In beiden Fächern reagieren die Hauptschüler mit einschneidenderen Leistungsveränderungen.

Es kann festgehalten werden, daß die experimentellen Bedingungen bei Entlaßschülern der Hauptschule neben häufigeren Veränderungen im Leistungsprozeß auch eindeutigere Leistungsschwankungen hervorrufen als bei Viertkläßlern. Besonders deutlich zeigt sich dieser Unterschied im Diktatversuch.

Neben Häufigkeit und Eindeutigkeit kennzeichnet vor allem die Richtung der Leistungsschwankungen unter dem Einfluß der experimentellen Bedingungen den Leistungsverlauf. Es fällt auf, daß der Vergleich KA -PA ein nahezu eindeutiges Verlaufsbild bietet. Abgesehen von einem Einzelabschnitt (9. Schuljahr, Rechnen, 2. Abschnitt) stellen sich alle Leistungsschwankungen als Leistungseinbußen heraus, die im 9. Schuljahr meist sehr eindeutig ausfallen (insgesamt acht Leistungseinbußen, davon sieben sehr deutlich). Wenn also im Vergleich KA - PA die Prüfungsbedingungen das Fehlerniveau beeinflussen (das ist bei 14 von 19 Differenzen geschehen), verschlechtern sich

Schüler fast immer in der nicht üblichen Prüfungsarbeit (PA) gegenüber der gewohnten Klassenarbeit (KA).

Ähnlich ausgeprägte Tendenzen treffen für die beiden anderen Vergleiche nicht zu. Es ist aber nicht von der Hand zu weisen, daß Schüler auch in den einzelnen Phasen der völlig entspannten Übungssituation (ÜA) oft besser abschneiden als im Verlauf der hoch belastenden Prüfungsarbeit (Ausnahme: 9. Schuljahr, Diktat, 2. Abschnitt), obwohl auf Grund von sieben signifikanten Leistungsunterschieden (bei 19 möglichen) deutlich wird, daß sich der Einfluß der experimentellen Bedingungen in diesem Bereich weniger eindeutig auswirkt.

Das Bild der Leistungsdifferenzen zwischen ÜA und KA wird von der eindeutigen Verbesserung der Schüler des 9. Schuljahres im Diktatversuch geprägt. Bis auf den dritten Einzelabschnitt wirkt die etwas angespannte Situation während der Klassenarbeit so leistungssteigernd in allen Phasen des Leistungsprozesses, daß sehr klare Verbesserungen gegenüber Ergebnissen in Übungsarbeiten zu registrieren sind (1%-Niveau). Der Tatbestand der Leistungsverbesserung in der Klassenarbeit trifft, wenn auch in weit schwächerem Maße, auch auf die Rechengruppe des 9. Schuljahres zu. Die beobachteten Verbesserungen (zweimal $p < .05$) sind aber schwach ausgeprägt, so daß sie sich im Gesamtergebnis nicht niederschlagen. Im 4. Schuljahr wirkt sich die Situation der Klassenarbeit in nicht bedeutsamem Maße auf die Einzelergebnisse des Leistungsverlaufs aus (einer Verbesserung steht eine Leistungsminderung entgegen). Auch das Ritual der Klassenarbeit veranlaßt Viertkläßler im Gegensatz zu Schülern der Entlaßklasse der Hauptschule nicht, ihre Leistungen im Verlauf einer solchen Arbeit zu steigern.

Zur Lokalisierung des Einflusses der experimentellen Bedingungen im Leistungsprozeß werden die Verlaufsbereiche I (erster und zweiter Einzelabschnitt) und II (vierter und fünfter Einzelabschnitt) analysiert.

Tabelle 6: Einfluß der Prüfungssituationen in den beiden Verlaufsbereichen (Prüfverfahren: FRIEDMAN-Test)

Schuljahr	Fach	I. Verlaufsbereich	II. Verlaufsbereich
4.	Rechnen	13.964	4.411
		<.01*	>.10
4.	Diktat	1.647	1.492
		>.10	>.10
9.	Rechnen	2.188	14.509
		>.10	<.01
9.	Diktat	17.320	11.092
		<.01	<.01

* Zufallswahrscheinlichkeit des Chi-Quadrat-Wertes

Auf Grund der Größe der Chi-Quadrat-Werte (Tabelle 6) läßt sich die Aussage vertreten, daß keiner der beiden Verlaufsbereiche den Einflüssen der Prüfungsbedingungen in erhöhtem Maße ausgesetzt ist (je zwei sehr signifikante Werte im ersten und zweiten Verlaufsbereich). Es fällt auf, daß Viertkläßler, wenn sie auf die durch Instruktionen erzeugten Arbeitsbedingungen reagieren, im Anfangsbereich mit entsprechenden Leistungen antworten. Die Arbeiten im Zahlenrechnen sind in etwa dem Schwierigkeitsanstieg der Aufgaben in den üblichen Klassenarbeiten angepaßt. Bemerkenswerterweise aber beziehen sich die Einflüsse der Prüfungssituationen bei Viertkläßlern gerade auf den leichteren Verlaufsbereich I, während im schwierigeren Bereich (Endleistung) Belastungsbeeinträchtigungen nicht

nachzuweisen sind. Im Rechenversuch des 9. Schuljahres hingegen reagieren Schüler gerade im Verlaufsbereich II (Endleistung) sehr empfindlich auf die unterschiedlichen Prüfungssituationen. Die Diktatergebnisse beider Schuljahre weisen solche schwerpunktartigen Wirkungen der experimentellen Bedingungen nicht aus.

Differenziertere Hinweise ergibt eine Überprüfung der Meßreihen mit dem WILCOXON-Test.

Tabelle 7: Leistungsdifferenzen in den beiden Verlaufsbereichen (Prüfverfahren: WILCOXON-Test)

Schuljahr	Fach	Vergleich	Verlaufsbereich I	Verlaufsbereich II
4.	Rechnen	ÜA - KA	-0,28	+0,36
			< .05*	< .05
4.	Rechnen	KA - PA	-0,42	+0,10
			< .01	> .10
4.	Rechnen	ÜA - PA	-0,70	+0,26
			< .01	> .05
4.	Diktat	ÜA - KA	+0,17	+0,01
			> .10	> .10
4.	Diktat	KA - PA	-0,46	-0,25
			> .05	> .10
4.	Diktat	ÜA - PA	-0,27	-0,24
			> .10	> .10
9.	Rechnen	ÜA - KA	+0,06	+0,02
			> .10	> .10
9.	Rechnen	KA - PA	+0,15	-0,95
			> .10	< .01
9.	Rechnen	ÜA - PA	+0,21	-0,93
			> .10	< .01
9.	Diktat	ÜA - KA	+1,57	+1,65
			< .01	< .01
9.	Diktat	KA - PA	-1,47	-1,12
			< .01	< .01
9.	Diktat	ÜA - PA	+0,10	+0,53
			> .10	> .05

* Zufallswahrscheinlichkeit der U-Werte

Rechenleistungen von Schülern des 4. Schuljahres sind besonders in der ersten Arbeitsphase (I. Verlaufsbereich) den leistungsmindernden Einflüssen der unterschiedlich belastenden Prüfungssituationen ausgesetzt. Durchgängig verschlechtern sich die Leistungen der Schüler, während in der Endphase der Einfluß der experimentellen Bedingungen sich entweder nicht leistungsändernd auswirkt oder sogar noch eine Leistungsverbesserung zuläßt (ÜA - KA), obwohl der Schwierigkeitsgrad der Aufgaben leicht ansteigt. Rechenleistungen im 9. Schuljahr leiden besonders in der Schlußphase (Endleistung) unter den nicht gewohnten, strengen Bedingungen der Prüfungsarbeit (PA). Die Belastungen der Klassenarbeit schlagen sich gegenüber Übungsarbeitsergebnissen nicht in Leistungsänderungen nieder.

Einflüsse der experimentellen Bedingungen wirken sich in den Diktatleistungen - wie schon der FRIEDMAN-Test nachweist - gleichmäßig aus. Die situativen Prüfungseinwirkungen sind im Diktatversuch des 4. Schuljahres so schwach, daß keinerlei gesicherter Einfluß nachweisbar ist. Im 9. Schuljahr schreiben Schüler unter Klassenarbeitsbedingungen eindeutig die besten Arbeiten. Der Einfluß der Prüfungsbedingungen wirkt sich allerdings in beiden Leistungsbereichen gleichmäßig stark (1%-Niveau) aus.

Zusammenfassend kann man konstatieren, daß Prüfungsbelastungen während des Arbeitsablaufes oft auch kleinere Leistungsintervalle in statistisch bedeutsamem Maße leistungsverändernd beeinflussen. Der Leistungsverlauf von Schülern des 9. Schuljahres ist nachhaltiger von Leistungsschwankungen geprägt, als es im Leistungsverlauf von Viertkläßlern nachweisbar ist. Die geprüften Schüler-

gruppen verschlechtern sich in den Teilabschnitten des Leistungsverlaufs zumeist unter dem Einfluß der ungewohnten Bedingungen der Prüfungsarbeit (PA), wenn man deren Ergebnisse mit den Daten der weniger belastenden Prüfungssituationen, besonders mit der Klassenarbeit, vergleicht. In den Ergebnissen des 9. Schuljahres deuten die Leistungsunterschiede in den Einzelabschnitten auf einen leistungsfördernden Einfluß der gewohnten, aber doch wohl etwas gespannten Klassenarbeitsatmosphäre hin (ausgeprägter: in den Diktatergebnissen des 9. Schuljahres).

Einflüsse von Prüfungssituationen können im Anfangs- und Endbereich des Leistungsprozesses Fehlerquoten signifikant ändern. Viertkläßler reagieren im Rechnen mit Leistungsminderungen im Anfangsbereich, wenn sie zunehmend strengen und harten Prüfungssituationen ausgesetzt werden, während Schüler des 9. Schuljahres im Leistungsprozeß unter dem Einfluß der ungewohnt harten Belastungsproben der Prüfungsarbeit (PA) zum Schluß ihrer Arbeiten einen Leistungsabfall erleiden.

Die experimentellen Bedingungen wirken sich in den Diktatversuchen beider Jahrgänge gleichmäßig in beiden Verlaufsbereichen aus. Schwerpunktartige Wirkungen können nicht in diesem Bereich studiert werden.

Art des Schulfaches bzw. die unterschiedliche Schwierigkeit der Rechenaufgaben führen nicht zu gehäuften, in nur einem Bereich lokalisierbaren Leistungsveränderungen (vgl. Zahlenrechnen: 4. und 9. Schuljahr).

Die Datenanalyse verbietet auch den Schluß, daß sich

gesteigerte Prüfungsbelastungen immer nur in der Endphase von Probearbeiten leistungsmindernd auswirken (wie SANDVOSS 1952 vermutet).

4. EINFLUß VON PERSÖNLICHKEITSMERKMALEN

Vorbemerkung

Auf Grund des großen Stichprobenumfanges (50 Schulklassen) und des zeitlich hoch bemessenen Unterrichtsaufwandes (sieben Unterrichtsstunden) war es im Rahmen dieser Arbeit kaum möglich, diejenigen Persönlichkeitsvariablen näher zu bestimmen, die über reine Leistungsfeststellungen hinausgehen und einen tieferen Einblick in die Motivationsstruktur des Probanden geben (bis auf den Versuch der Einstellungsmessung). So wurden geplante Untersuchungen zur Ermittlung der Schulbegabung, der Ängstlichkeit und der Leistungsmotivation abgebrochen, da die mitarbeitenden Lehrer trotz Genehmigung der Untersuchung durch das Kultusministerium von NRW einerseits den beträchtlichen Zeitaufwand des Unterrichtsversuchs bemängelten und nur selten bereit waren, die erforderliche Zeit für die Aufnahme von Persönlichkeitsdaten bereitzustellen, und andererseits auf den für sie angeblich wenig informativen Charakter der angestrebten Persönlichkeitsuntersuchungen hinwiesen. Außerdem befürchteten sie Interventionen von Seiten der Eltern. So konnten nur Schulleistungstests durchgeführt werden, deren Ergebnisse mit den letzten Zeugniszensuren im betreffenden Fach kombiniert wurden, um die einzelne Klasse bezüglich der entsprechenden Fachleistung in eine Rangordnung zu bringen. Das Zusammenwirken von experimentellen Bedingungen und anderen persönlichkeits-

spezifischen Faktoren als Quelle der Fehlervarianz konnte leider nicht im wünschenswerten Umfang studiert werden.

4.1. Prüfungssituation und Fachleistung

Da das Aufgabenmaterial ausnahmslos schulspezifischen Charakter hatte, lag es nahe, auch an den Einfluß von Faktoren des fachlichen Leistungsvermögens auf Eindeutigkeit und Richtung von Leistungsschwankungen in den Prüfungssituationen zu denken. Ausschlaggebend für die Erstellung der Rangordnung war die letzte Zeugniszensur im betreffenden Fach. Die Rangordnung der Schüler mit der gleichen Zensur ergab sich aus den Werten der Schulleistungstests (9. Schuljahr) oder der durchschnittlichen Fehlerzahl aller Arbeiten, weil ein geeigneter Rechentest im 4. Schuljahr nicht zur Verfügung stand. Es mußte gewährleistet sein, daß jede Klasse der Anzahl von Schülern entsprechend in den Leistungsgruppen vertreten war. Alle Schulklassen wurden auf Grund der Rangordnungen in Fünftel aufgeteilt. Das beste Fünftel wurde der stärksten Leistungsgruppe (I) zugeteilt, das mittlere Fünftel auch der mittleren Leistungsgruppe (II), das schwächste Fünftel jeweils auch der schwächsten Leistungsgruppe (III) zugerechnet.

Tabelle 8 gibt Auskunft, inwieweit innerhalb der Leistungsgruppen Prüfungseinflüsse wirksam sind.

Tabelle 8: Einwirkungen der experimentellen Bedingungen in den Gesamtergebnissen der Leistungsgruppen (Prüfverfahren: FRIEDMAN-Test)

Schuljahr	Fach	Leist.Gr.	Vpn	Chi-Quadrat-Wert
4.	Rechnen	I	35	3.186 >.10*
4.	Rechnen	II	35	0,871 >.50
4.	Rechnen	III	35	3.443 >.10
4.	Diktat	I	48	0,219 >.50
4.	Diktat	II	48	0,448 >.50
4.	Diktat	III	48	4,823 >.05
9.	Rechnen	I	32	1.750 >.10
9.	Rechnen	II	32	2.919 >.10
9.	Rechnen	III	32	4.703 >.05
9.	Diktat	I	42	14.393 <.01
9.	Diktat	II	42	6.298 <.05
9.	Diktat	III	42	15.083 <.01

* Zufallswahrscheinlichkeit der Chi-Quadrat-Werte

Belastungseinflüsse durch Prüfungsbedingungen lassen sich nur in allen Leistungsgruppen des Diktatversuches im 9. Schuljahr einwandfrei nachweisen. Alle anderen Werte sprechen dafür, daß bedeutsame Unterschiede in diesem Zu-

sammenhang nicht vorliegen. Es ist aber auf Grund der leicht erhöhten Chi-Quadrat-Werte der schwächsten Leistungsgruppe nicht ganz von der Hand zu weisen, daß sich durch entsprechende Instruktionen gesteigerte psychische Belastung auch schon bei der Bearbeitung von gewohnten schulischen Inhalten gerade bei leistungsschwachen Schülern auswirkt. Allerdings ist Zurückhaltung bei der Deutung solcher Tendenzen geboten, da sie nicht statistisch gesichert sind.

Erst Prüfung der Daten mit dem WILCOXON-Test und evtl. Anschlußexperimente werden diese Frage verbindlicher beantworten können.

In Tabelle 9 sind die Leistungsunterschiede KA - PA und ÜA - PA wiedergegeben; die Differenzen ÜA - KA sind ausgespart worden, weil sich in diesem Vergleich keinerlei Hinweise auf die Einflüsse der Fachleistung im Fehlerverhältnis finden lassen.

Tabelle 9: Einflüsse der experimentellen Bedingungen auf die Leistungsdifferenzen KA - PA und ÜA - PA in den Leistungsgruppen (Gesamtergebnis) (Prüfverfahren: WILCOXON-Test)

Schuljahr	Fach	Leist.Gr.	Vpn	Vergleich KA - PA	Vergleich ÜA - PA
4.	Rechnen	I	35	-0,80	-1,00
				<.05*	<.05
4.	Rechnen	II	35	+0,03	-0,43
				>.50	>.10
4.	Rechnen	III	35	-0,66	-0,77
				<.05	>.05
4.	Diktat	I	48	-0,17	-0,23
				>.10	>.50
4.	Diktat	II	48	-0,88	-0,23
				>.05	>.50
4.	Diktat	III	48	-0,73	-0,87
				>.10	>.05
9.	Rechnen	I	32	-0,78	-1,19
				>.05	>.05
9.	Rechnen	II	32	-1,64	-0,99
				<.01	>.05
9.	Rechnen	III	32	-1,42	-1,79
				<.01	<.05
9.	Diktat	I	42	-3,43	-0,43
				<.01	>.10
9.	Diktat	II	42	-2,40	+0,95
				<.01	>.10
9.	Diktat	III	42	-4,89	+0,50
				<.01	>.50

* Zufallswahrscheinlichkeit der U-Werte im Wilcoxon-Test

In den Diktatergebnissen des 4. Schuljahres wirkt sich die Leistungsstärke der Schüler nicht auf das Fehlerverhältnis aus. Die Fachleistung, so wie sie in dieser Arbeit formuliert wird, scheidet in diesem Falle als Ursache der Varianz aus.

Im Zahlenrechnen hingegen scheint die Variable Fachleistung die Größe der Leistungsdifferenzen zu beeinflussen. Gerade in der besten und schwächsten Leistungsgruppe des 4. Schuljahres wirken sich die ungewohnten Bedingungen der Prüfungsarbeit (PA) gegenüber den anderen Probesituationen leistungshemmend aus, während die mittlere Leistungsgruppe solchen Einflüssen gegenüber resistent bleibt und mit etwa gleichbleibenden Ergebnissen antwortet.

Ein ähnliches Ergebnis bietet die statistische Überprüfung der signifikanten Differenzen im Diktatversuch des 9. Schuljahres mit dem U-Test von MANN-WHITNEY (1947)[1]. Danach reagiert die Mittelgruppe mit der kleinsten negativen Differenz auf den Wechsel von Bedingungen der Klassenarbeit zum ungewohnten, strengen 'Ritual' der Prüfungsarbeit (PA), während gerade die leistungsschwachen Schüler noch deutlichere Leistungseinbußen beim Wechsel KA - PA erleiden als ihre leistungsstärksten Kameraden.

Im Rechenversuch des 9. Schuljahres lassen die signifikanten Prüfwerte (KA - PA; ÜA - PA) vermuten, daß ebenfalls die leistungsschwächeren Schüler unter den verschärften, ungewohnten Bedingungen der Prüfungsarbeit schlechter abschneiden als die leistungsstarken Schüler, deren Ergebnisse durch die experimentellen Bedingungen nur unwesentlich verändert werden. Der Eindruck, daß Schüler mit etwa mittlerer Leistungsstärke weniger anfällig für verschärfte Prüfungsbedingungen sind, wird durch ein überraschendes Einzelergebnis bestätigt, das erst in diesem Zusammenhang eine Bedeutung erhält. So weist eine Analyse des Leistungsverlaufs im Zahlenrechnen der mittleren Leistungsgruppe des 4. Schul-

[1] Berechnung nach G.A. LIENERT, Verteilungsfreie Methoden in der Biostatistik, 1962

jahres im letzten Einzelabschnitt eine signifikante Steigerung in der ungewohnten, harten Prüfungsarbeit gegenüber Klassenarbeitsergebnissen auf, ein Sachverhalt, der in den Ergebnissen der Extremgruppen keine Parallele findet.

Als Schlußfolgerung aus diesen Daten kann festgehalten werden, daß die Qualität der Fachleistung als Differenzierungsvariable bezüglich der Wirkung der experimentellen Bedingungen im Rechtschreiben für Viertkläßler und für bedeutsame Leistungsänderungen beim Wechsel von der Übungssituation zu den Bedingungen der Klassenarbeit (ÜA - KA) im Gesamtversuch nicht verantwortlich ist. Im Zahlenrechnen sowie im Diktatversuch des 9. Schuljahres können Zusammenhänge zwischen Fachleistung und Erhöhung der Fehlerquoten beim Wechsel von den schulüblichen Probesituationen (ÜA, KA) zur ungewohnten Prüfungsarbeit nicht von der Hand gewiesen werden. Die mittlere Leistungsgruppe scheint in diesem Zusammenhang weniger deutlich mit Leistungsänderungen zu reagieren als die Extremgruppen, von denen gerade die leistungsschwache Gruppe einen erhöhten Leistungsabfall in der Prüfungsarbeit erleidet, wenn man die Ergebnisse aus den üblichen Arbeitssituationen zum Vergleich heranzieht.

4.2. Prüfungssituation und Geschlecht

Grundsätzlich wurden nur Schulen berücksichtigt, in denen gemischte Klassen unterrichtet wurden. Deshalb werden auch beide Geschlechter zur Befundanalyse herangezogen. Dies entspricht von wenigen Ausnahmen abgesehen auch der heutigen Schulpraxis. Die folgende Analyse soll nur insoweit dargelegt werden, wie sie dazu beiträgt, Ergebnisse, die vom Trend der bisher dargestellten Befunde abweichen, zu verdeutlichen.

Die Daten der Tabelle 10 lassen erkennen, daß insgesamt nur wenig präzise Angaben über geschlechtsspezifische Unterschiede in diesem Unterrichtsversuch herausgearbeitet werden können.

Tabelle 10: Einfluß der experimentellen Bedingungen im Gesamtbild von Jungen und Mädchen (Prüfverfahren: FRIEDMAN-TEST)

Schuljahr	Fach	Geschlecht	Vpn	Chi-Quadrat-Wert
4.	Rechnen	Jungen	106	3.995 >.10*
4.	Rechnen	Mädchen	73	2.308 >.10
4.	Diktat	Jungen	111	3.992 >.10
4.	Diktat	Mädchen	117	0.594 >.50
9.	Rechnen	Jungen	88	5.528 >.05
9.	Rechnen	Mädchen	74	9.270 <.01
9.	Diktat	Jungen	90	16.017 <.01
9.	Diktat	Mädchen	93	17.145 <.01

* Zufallswahrscheinlichkeit der Chi-Quadrat-Werte

Nach den vorliegenden Ergebnissen besteht nur in den Rechenleistungen des 9. Schuljahres ein gesicherter Unterschied im Wirkungsgrad der Prüfungsbedingungen, denn während die Leistungen der Schülerinnen einen solchen Prüfungseinfluß eindeutig bestätigen ($p < .01$), bleiben die Leistungen der Jungen in diesem Bereich relativ unbeeindruckt von der Wirkung der unterschiedlichen Prüfungssituationen ($p > .05$).

Im einzelnen (WILCOXON-Überprüfung) reagieren die Mädchen beider Schuljahre in Arbeiten mit reinem Zahlenrechnen beim Wechsel von der schulüblichen Arbeitssituation (ÜA,KA) zur ungewohnten Prüfungsarbeit (PA) mit schärferen Leistungsminderungen als die Jungen (Tabelle 11).

Tabelle 11: Einfluß der experimentellen Bedingungen auf die Gesamtergebnisse des Zahlenrechnens von Mädchen und Jungen (Prüfverfahren: WILCOXON-Test)

Schuljahr	Fach	Geschlecht	Vergleich KA - PA	Vergleich ÜA - PA
4.	Rechnen	Jungen	-0,42	-0,06
			<.05*	>.50
4.	Rechnen	Mädchen	-0,66	-0,96
			<.01	<.01
9.	Rechnen	Jungen	-1,02	-0,78
			<.05	>.05
9.	Rechnen	Mädchen	-1,58	-1,14
			<.01	<.05

* Zufallswahrscheinlichkeit der U-Werte im Wilcoxon-Test

Die Überprüfung der Differenzen mit dem U-Test von MANN-WHITNEY bestätigt die Erkenntnis, daß Mädchen empfindlicher in leistungsmindernder Richtung auf erhöhten situativen Druck in Prüfungssituationen reagieren. Mit Blick auf die Diktatergebnisse läßt sich dieser Befund allerdings nur für das Zahlenrechnen bestätigen.

Eine Analyse des Leistungsverlaufs ergibt keine Einzelergebnisse, die geschlechtsspezifische Unterschiede im Leistungsverhalten hinreichend sichern können. Diese Einzelbefunde müssen vor dem Hintergrund durchgängig festgestellter Übereinstimmungen im Leistungsverhalten

von Jungen und Mädchen gesehen werden. Mit Ausnahme der oben besprochenen geschlechtsspezifischen Reaktionen, die sich aber auf die Nachhaltigkeit eines Einflusses und nicht auf die Richtung beziehen, können die unter 3.1. und 3.2. genannten Befunde auch für Jungen und Mädchen geltend gemacht werden.

4.3. Prüfungssituation und Einstellung zum Schulfach

Es soll der Frage nachgegangen werden, ob fachinteressierte Schüler, die das Fach und die damit verbundenen Aufgaben gern erledigen, sich im Leistungsverhalten unterscheiden vom in dieser Hinsicht nicht auffälligen Durchschnitt oder von Schülern, die dem Fach und den Aufgaben ablehnend gegenüberstehen. Die Auswertung der Unterlagen ergibt sehr unterschiedliche Beliebtheitsgrade für die beiden betroffenen Schulfächer. Während es im Rechnen nicht gelang, eine ausreichend umfangreiche Stichprobe fachablehnender Schüler zusammenzustellen (es hätte einer beträchtlichen Erhöhung der Gesamtstichprobe bedurft), ließen entsprechende Angaben im Rechtschreibversuch des 9. Schuljahres einen solchen Vergleich zu (9. Schuljahr: 33 zustimmende, interessierte Schüler, 36 fachablehnende Schüler).

Insgesamt lassen sich auch im Rechtschreibversuch des 9. Schuljahres nur wenige Anhaltspunkte für einen Zusammenhang von Fachinteresse und Abschneiden unter dem Druck unterschiedlich belastender Prüfungssituationen feststellen. Interessierte Schüler verschlechtern sich in der ungewohnten Prüfungsarbeit (PA), ein Effekt, der bei fachablehnenden Schülern weder im Gesamtergebnis noch in Einzelabschnitten des Leistungsverlaufs nachzuweisen ist. Die Datenbasis wird in Anschlußexperimenten weiter

ausgebaut werden, um evt. vorhandene Effekte dieser Variablen eindeutiger zu umreißen.

4.4. Prüfungssituation und Zeichensetzung

Im 4. Schuljahr werden Satzzeichen mitdiktiert, im 9. Schuljahr müssen Satzzeichen herausgefunden und eingetragen werden. Man darf vermuten, daß Satzzeichen sowohl im 4. Schuljahr (weil sich Fehler dieser Art zumeist nicht als Fehlleistungen im Gesamtergebnis niederschlagen) als auch im 9. Schuljahr (weil Fehler in der Zeichensetzung gar nicht oder nur mit einem halben Fehler gewertet werden) nicht soviel Beachtung finden wie der diktierte Text. Da zudem stark belastende, wenn auch nicht ungewohnte Durchführungsbestimmungen des Diktatversuchs (die Schüler hören den Text nur insgesamt zweimal) das Hauptaugenmerk der Schüler auf das Behalten des Textes lenken, ist zu erwarten, daß analog zur Bedeutung eines möglichst guten Abschneidens, das durch vollständiges und fehlerfreies Reproduzieren und Niederschreiben einer einmal vorgelesenen 'Sinneinheit' gesichert ist, sich die Zahl der Z-Fehler erhöht, weil die Satzzeichen als unwichtiges Anhängsel betrachtet werden.

Tabelle 12: Einfluß der experimentellen Bedingungen auf Wort- und Zeichensetzungsfehlerverteilung (W- und Z-Fehler) (Prüfverfahren: FRIEDMAN-Test)

Schuljahr	Fach	Fehlerart	Chi-Quadrat-Werte
4.	Diktat	W	0,712
			>.50*
4.	Diktat	Z	2.028
			>.10
9.	Diktat	W	24.734
			≤.01
9.	Diktat	Z	12.167
			≤.01

* Zufallswahrscheinlichkeit der Chi-Quadrat-Werte

Die Prüfung der Rangsummenunterschiede läßt keine Aussage über die Gültigkeit der oben besprochenen Hypothese zu.

Tabelle 13: Einfluß der experimentellen Bedingungen auf die Verteilungen der W- und Z-Fehler (Prüfverfahren: WILCOXON-Test)

Schuljahr	Fehlerart	Vergleich ÜA - KA	Vergleich KA - PA	Vergleich ÜA - PA
4.	W-Fehler	-0,18	-0,38	-0,56
		>.10*	<.05	>.05
4.	Z-Fehler	-0,03	-0,13	-0,16
		>.10	<.05	<.05
9.	W-Fehler	+3,00	-1,85	+1,15
		<.01	<.01	<.05
9.	Z-Fehler	+0,52	-1,09	-0,57
		>.05	<.01	<.05

* Zufallswahrscheinlichkeit der U-Werte im Wilcoxon-Test

Da der WILCOXON-Test auf Rangebene prüft, können auch vergleichsweise kleine Differenzen in der zentralen Tendenz, wie sie bei Z-Fehlern im 4. Schuljahr zu beobachten sind, statistisch gesichert sein.

Anhand der Daten der Tabelle 13 ist zu ersehen, daß unterschiedlich belastende Prüfungssituationen die Höhe der Z-Fehler stärker beeinflussen kann als die der W-Fehler. Zwar ist der Grad des leistungsmindernden Einflusses der Prüfungsarbeit im Vergleich KA - PA etwa gleich bei beiden Fehlerarten, aber es ist auch unabweisbar, daß sich der Druck der ungewohnten Prüfungsarbeit weitaus leistungshemmender in der Z-Fehlerverteilung auswirkt als bei den W-Fehlern, wenn man Ergebnisse in der entspannten Übungssituation als Vergleichsmaßstab heran-

zieht. Sogar im Falle einer signifikanten Abnahme der W-Fehler (9. Schuljahr: ÜA - PA) schlägt sich der Einfluß der Prüfungsarbeit im Ergebnis der Z-Fehler als Leistungseinbuße nieder. Ein anderer Sachverhalt unterstützt diesen Befund. Schüler des 9. Schuljahres verbessern sich in Klassenarbeiten gerade im Bereich der W-Fehler sehr eindeutig (ÜA - KA), haben aber ansonsten in Klassenarbeiten ebensoviele Z-Fehler aufzuweisen wie in Übungsarbeiten.

Gesteigerte Belastungen (ÜA-PA, KA-PA) beeinflussen also im Sinne der oben dargestellten Hypothese die Verteilung der Zeichensetzungsfehler nachhaltiger in leistungsmindernder Richtung als die W-Fehler. Auch in den Bereichen, in denen die gesteigerte Bedeutung einer Prüfungssituation sogar eine Verbesserung der W-Fehlerquote bedingt, (9. Schuljahr, ÜA-KA, ÜA-PA), ändert sich die Z-Fehlerquote nicht oder auch nur in Richtung der Fehlererhöhung.

Es ist demnach - wenn auch bei noch nicht absolut lückenloser Beweisführung - nicht von der Hand zu weisen, daß gesteigerter Leistungsdruck in Prüfungssituationen den Blick auf die als wichtig erachteten und daher beachteten Teile der Aufgabenerledigung einengt und weniger bedeutungsvolle Aufgabenteile der Aufmerksamkeit des Schülers entzieht.

5. ANSÄTZE EINER THEORIEBILDUNG

Die ausführliche Befundanalyse in diesem Experiment läßt die Mehrdimensionalität des Problems erkennen, das der Schule mit dem Phänomen 'Prüfung, Leistungskontrolle' gestellt ist. Unterschiedlich belastende Prüfungssituationen wirken sich in der Gesamtstichprobe der Schüler auch unterschiedlich auf das Leistungsverhalten aus; es lassen sich, sofern Struktur des Prüfmaterials, Art der Prüfungssituation und Alter der Schüler berücksichtigt werden, sowohl Leistungsverschlechterungen als auch - allerdings weniger häufig - Verbesserungen nachweisen. Oft wirkt sich ein Wechsel der Instruktion nicht leistungsändernd aus.

Erklärungen für diese z.T. unterschiedlichen Leistungsbilder können nur dann befriedigen, wenn Forschungsergebnisse sozialpsychologischer Art und gesicherte Erkenntnisse bezüglich der Entwicklung des Leistungsverhaltens und der sozialen Funktion der Schule in der modernen Welt als Erklärungsgrundlage miteinbezogen werden. Es ist eben nicht möglich, Leistungsphänomene in einem Feldexperiment zu deuten, wenn nicht die Entwicklung, die "Organisation des gesamten Feldes" (WERTHEIMER 1945, 1957, in ähnlichem Sinne LEWIN 1963 und WINNEFELD 1957) ins Auge gefaßt wird.

Grundvoraussetzung für das reibungslose Funktionieren des Leistungswettbewerbs in der Schule ist die Entwicklung der Leistungsmotivation beim Kind. HECKHAUSEN (1969) definiert Leistungsmotivation "als das Bestreben, die

eigene Tüchtigkeit in allen jenen Tätigkeiten zu steigern oder möglichst hoch zu halten, in denen man einen Gütemaßstab für verbindlich hält und dessen Ausführung deshalb gelingen oder mißlingen kann." Leistungsmotivation ist nach GOTTSCHALCH, NEUMANN-SCHÖNWETTER, SOUKUP (1971) eine Konsequenz der elterlichen Erziehung und des häuslichen Anregungsmilieus. HECKHAUSEN, ROELOFSEN (1962) weisen in ihren Experimenten auf die Ausbildung des Leistungsverhaltens schon vor dem Schuleintritt hin.

Die Schule als weiterer Erziehungsträger baut die schon gefestigte Leistungsorientierung des Schulanfängers aus, indem sie mit einer Anzahl von Kontrolltechniken im permanenten Wettbewerb das System gesellschaftlich relevanter Leistungsideale einübt und ausdehnt. Die Internalisierung dieses Systems ist ein Lernprozeß, der mit fortschreitender Sozialisierung zu einer weitgehenden Identifizierung des Individuums mit dem Normensystem der Umwelt führt. Inhaltlich lernt der Schüler im Wettbewerb mit anderen sozial gebilligte Verhaltensweisen wie: Erster, Bester sein, sich nicht blamieren, nicht versagen, Kontrolle ohne schwache Zensur überstehen, Versetzung erreichen, günstige Startposition für die Berufsausbildung in Form 'guter' Zensuren erreichen. v. HENTIG (1971) beschreibt die Vorgehensweisen, deren sich die Schule bei der Einübung der Leistungsideale bedient: "Vom ersten Schultag an werden die Kinder durch Belohnung und Strafe, durch Nachsitzen als 'Sühnemaßnahme' und ständige Denunziation bei den Eltern, durch Bürokratisierung der Beurteilungsprozedur und durch öffentliches Lob und öffentliche Bloßstellung dazu abgerichtet, ihre Tätigkeit an der Tätigkeit der anderen und der in ihr zum Ausdruck kommenden abstrakten Erwartung zu messen." Der Lehrer, Träger der Normen der Umwelt (meist der Mittelstandsnormen) übernimmt die Auf-

gabe, diese gesellschaftlich gebilligten Leistungsnormen mittels eines ihm mehr oder weniger aufgezwungenen Wettbewerbs- und Kontrollsystems einzuüben. Einschlägige Anweisungen zur Leistungsfeststellung im schulischen Unterricht sind in den Richtlinien des Landes NRW niedergelegt. Erst auf dem Hintergrund dieser Erörterungen gewinnen die Ergebnisse der Befundanalyse Bedeutung.

Voruntersuchungen im Hinblick auf den Diktierversuch und die Kontrollergebnisse im Rechenversuch bestätigen die Vermutung, daß bei Identität der äußeren Bedingungen und vor allem der Instruktion die beobachteten Fehlerverhältnisse zwischen den einzelnen Arbeiten (erste und zweite Arbeit, zweite und dritte Arbeit, erste und dritte Arbeit) im Diktat nur unwesentlich voneinander abweichen. Im Rechenversuch verringert sich die Fehlerquote von der ersten zur dritten Arbeit relativ gleichmäßig (Wiederholungseffekte !). Einschneidende Änderungen dieser Fehlerverhältnisse treten erst im Hauptversuch auf. Entscheidende Abweichung im Versuchsplan ist die unterschiedliche Arbeitsinstruktion. Man darf wohl vermuten, daß das beobachtete Leistungsverhalten im Hauptversuch nach weitgehender Ausschaltung von Störvariablen in erster Linie durch unterschiedlich anregende Instruktionen und Arbeitsmittel (Übungshefte, Klassenarbeitshefte, vom Lehrer ausgehändigte DIN-A-4-Blätter) bestimmt wird. Der unterschiedliche Anregungsgrad der Instruktionen verhindert, daß die Schüler die Prüfungssituationen hinsichtlich des Leistungsanspruchs und Belastungsgrades als identisch erleben. Genauere Maße, die eine differenzierte Aussage zum Erleben von Prüfungssituationen sichern können, sind nicht vorweis-

bar, da u.a. präzise Angaben über aktuelle Leistungsangst und aktualisierte Motivation nicht zu erhalten sind.

Als erster Ansatz in einer Theorie der schulischen Leistung in Prüfungssituationen kann auf Grund der vorherigen Angaben festgehalten werden:

Schulerfahrene Viertkläßler und Entlaßschüler der Hauptschule reagieren, wie sie es gelernt haben, empfindsam auf schulische Leistungsforderungen, die der Lehrer formuliert. Lehrer beeinflussen das Leistungsverhalten der Schüler durch den Inhalt ihrer Instruktionen und Bereitstellung bedeutungsvoller Arbeitsmittel. Sie induzieren und reduzieren Stress (seelische Belastung) durch verbales Eingreifen.

Auch andere Befunde unterstreichen die verantwortungsvolle Rolle des Lehrers, wenn er durch Stressinduzierung Leistungsforderungen gegenüber seinen Schülern durchsetzen will.

Sieht man von den Diktatergebnissen des 9. Schuljahres und einigen, aber nicht überzubewertenden Leistungsverbesserungen in den Einzelabschnitten des Leistungsverlaufs im Zahlenrechnen des 9. Schuljahres ab, so ändert sich das Fehlerniveau im Vergleich ÜA - KA nur unwesentlich, obwohl der Hinweis auf die Bedeutung der Klassenarbeit, auf erhöhte Anstrengung zu einem verstärkten Bemühen um saubere strukturierte Darstellung (Drill!) führt. Aufforderung zu erhöhter Anstrengung ruft in dieser Situation keine negative Leistungsänderung gegenüber Ergebnissen in der entspannten Übungssituation (trotz Lehrerkontrolle, Zensierung, Elternbenachrichtigung, Aufbewahren der Arbeit) hervor. Es ist erstaunlich, daß die negative Reaktion auf die schärfere,

mehr Anstrengung fordernde Instruktion gerade dann ausfällt, wenn die Belastung im Rahmen der gewohnten Unterrichtsarbeit erhöht wird, indem die schon oft erlebten Instruktionen der Klassenarbeiten geboten werden. Die Schüler haben die Leistungsfeststellungen durch Klassenarbeiten vom ersten Schuljahr an erfahren und empfinden diese Situation zu diesem Zeitpunkt (nach dreieinhalb Jahren Unterricht) als durchschaubar und gewohnt, als notwendig und nicht umgehbar, als selbstverständlich. Auf Grund der Befunde darf man annehmen, daß ein Leistungsoptimum durch gewohnte, oft erlebte Aktivierung (Bedeutung der Klassenarbeit) erreicht wird (9. Schuljahr, Diktat) oder aber zumindest nicht überschritten wird (4. Schuljahr, 9. Schuljahr, Zahlenrechnen) und Leistungsminderungen nicht eintreten, wenn man die Ergebnisse anderer Probesituationen (ÜA, PA) als Vergleichsbasis heranzieht.

Steigert allerdings der Lehrer das Ausmaß der seelischen Belastung durch Aktualisierung von als existentiell wichtig erachteten Zielen (Versetzung, Übergangs-, Abschlußzeugnis) in seiner Instruktion, so steigt die Fehlerrate bei allen Schülern stark an (Vergleich KA - PA). Als einmalig wichtig ausgewiesene Leistungsnachweise in Form alles entscheidender Abschlußarbeiten sind den Schülern unbekannt, weil von ihnen nur studienbegleitende Klassenarbeiten verlangt werden. Leistungsverbesserungen können durch beliebige Erhöhung der seelischen Belastung also nicht erzwungen werden.

Der Lehrer sollte sich auch darüber im klaren sein, daß eine Belastungssteigerung über den gewohnten Rahmen hinaus eine Gruppe von Schülern besonders hart trifft. Leistungsschwache verschlechtern sich gerade in der ungewohnt belastenden Prüfungsarbeit mehr als alle anderen Schüler.

Überhöhter Druck führt also gerade dort, wo Lehrer sich Leistungssteigerungen erhoffen, zum rapiden Leistungsabfall. Als Bestätigung können zudem die Ergebnisse zahlreicher Untersuchungen herangezogen werden, in denen nachgewiesen ist (COX 1962, SASSENRATH 1963, STEVENSON, ODOM 1965, FELDHUSEN und KLAUSMEIER 1962), daß Ängstlichkeit, also in erster Linie Leistungsangst, wie sie in den gängigen Tests gemessen wird, und verschiedene Variable der Schulleistung negativ korrelieren. Dies bedeutet, daß in der Regel die Leistungsschwachen ängstlicher auf Leistungssituationen reagieren als die übrigen Gruppen und auf erhöhten Leistungsdruck mit starker leistungsmindernder Angstentwicklung antworten.

Insgesamt scheinen die vorliegenden Befunde den Gesetzmäßigkeiten des 'YERKES-DODSON-Gesetzes' (YERKES-DODSON 1908) zu folgen. Das YERKES-DODSON-Gesetz sagt im wesentlichen aus, daß die Beziehung zwischen der als Aktivierung, Anstrengung oder Belastung aufgefaßten Angst und der Leistung eine kurvilineare Funktion darstellt. Ein niedriges Aktivierungsniveau erschwert eine Leistung. Ein sehr hohes Aktivierungsniveau beeinflußt den Leistungsprozeß dahingehend, daß die Leistung absinkt, oft noch tiefer als bei niedrigem Aktivierungsniveau. Das Aktivierungsniveau, das zum Leistungsoptimum führt, liegt irgendwo im mittleren Bereich der Aktivierungsintensität. Diese in Versuchen mit Mäusen gesicherten Befunde lassen nach METZGER (1962) "eine Verwandtschaft mit den Ergebnissen über die Beeinträchtigung (und gelegentliche Förderung) von Denkvorgängen durch zusätzliche Belastung" vermuten.

HEBB (1955), ein Neurologe, sieht ebenfalls einen mittleren Grad der Aktivierung des Großhirns, bei dem alle geistigen Funktionen, einschließlich der Beherrschung und Steuerung

der eigenen Bewegungen, die günstigsten Bedingungen herrschen, als optimalen Leistungshintergrund an (zit. nach METZGER 1962).

Folgende für die theoretische Einbettung wichtige Grundvorstellung läßt sich unter Einbeziehung der Ergebnisse von SANDVOSS (1952) auf Grund der vorliegenden Befunde vertreten:

Zu Anfang ihrer Schullaufbahn empfinden Schüler auch schon Klassenarbeitssituationen als ungewohnt, besonders gefährlich, weil sie nicht genau wissen, was sie erwartet. Diese Prüfungssituation setzt sie so in Aufregung, daß ein Leistungsabfall eintritt gegenüber den Ergebnissen, die sie in der Übungssituation vorzuweisen hatten (SANDVOSS 1952). Es scheint, daß das für ein Leistungsoptimum wichtige mittlere Aktivierungsniveau bereits überschritten ist (YERKES-DODSON-MODELL). Die Viertkläßler hingegen haben gelernt, Klassenarbeiten einzuschätzen und mit ihnen zu leben. Die entspannte Übungsatmosphäre und die Klassenarbeitssituation bringen in etwa identische Fehlerraten hervor. Eine Erklärung für diesen Befund bietet die schulische Alltagserfahrung, daß für Schüler des 4. Schuljahres jede schulische Leistungssituation eine Aufforderung zur Anstrengung und Aktivierung in sich trägt. Im 9. Schuljahr dagegen scheint die Differenzierung zwischen relativ unbedeutender Übungsarbeit und wichtiger Klassenarbeit soweit bewußt zu sein, daß sich die Schüler erst in der brenzlicheren Situation der Klassenarbeit so anstrengen, daß sie ihre volle Leistung erreichen.

Viertkläßler und Entlaßschüler der Hauptschule reagieren wie Schulneulinge bei Klassenarbeiten (SANDVOSS) in der

ungewohnten, als einmalig wichtig herausgestellten Prüfungsarbeit mit eindeutigen Leistungsminderungen, wenn man die Ergebnisse in den gewohnten Arbeitssituationen (ÜA und KA) als Vergleichsmaßstab heranzieht. METZGER (1962) erklärt diesen Befund recht anschaulich: "Zusätzlicher Druck kann den Verschlafenen aufwecken, den Trägen ermuntern und dadurch fördern. Er wird aber denjenigen, der sich schon voll einsetzt, in einen Erregungszustand versetzen, bei dem er nicht mehr sein Bestes bietet". Autoren wie EYSENCK (1957) und BROADHURST (1959) weisen darauf hin, daß die Auswirkung des Aktivierungsgrades auf das Leistungsverhalten nicht zuletzt auch von der Schwierigkeit, der Lösbarkeit, der Komplexität des Aufgabenmaterials abhängig ist, so daß der Geltungsbereich der vorliegenden Befunde nicht überschätzt werden darf. Im vorliegenden Versuch sollte in erster Linie die unterschiedlich anregende Prüfungssituation als belastender Faktor geprüft werden, ohne daß Wechselwirkungen mit anderen situativen Variablen zu analysieren waren (Zeitdruck, Anreizgrad, Fremdheit, Schwierigkeit des Prüfmaterials).

6. ZUSAMMENFASSUNG

In der vorliegenden Arbeit wird das Leistungsverhalten von Viertkläßlern und Entlaßschülern von Hauptschulen im Zahlenrechnen und Rechtschreiben unter variierten Prüfungsbedingungen - in Anlehnung an ein Experiment von SANDVOSS (1952) - untersucht. Jeder Schüler unterzieht sich allen geplanten Prüfungen.

Schulklassen werden nach vorher festgelegten Merkmalen ausgewählt, um den Einfluß intervenierender Variablen so weit wie möglich auszuschalten (vornehmlich Frontal-

unterricht, Schülerfrequenz, Lage des Schulortes, Diktierstil, Koedukation). In Vorversuchen kristallisieren sich Struktur und Umfang des Prüfmaterials heraus. Außerdem werden Instruktionen und Verfahrensweisen festgelegt, um möglichst große Identität der Versuchsbedingungen zu garantieren.

Es kommt in diesem Feldexperiment darauf an, die Effekte schulnaher Prüfungssituationen mit möglichst schulüblichem Prüfungsstoff zu untersuchen. Die unterrichtsspezifische Ausgestaltung des Versuchs ist dadurch gewährleistet, daß sich das gesamte Experiment nahtlos in den normalen Unterrichtsverlauf einfügt, ohne daß die Schüler etwas vom Versuchscharakter der Situation vernehmen. Die Ausschaltung lehrerpersönlichkeitsspezifischer Faktoren sichert ein unterrichtserfahrener Versuchsleiter, der während der gesamten Experimentierphase den Unterricht im entsprechenden Fach übernimmt und als vollwertiger Lehrer anerkannt wird. Das Experiment erstreckt sich - die Aufnahme der Persönlichkeitsdaten nicht mitgerechnet - über sieben Unterrichtsstunden (7 mal 45 Minuten), wobei die Eingewöhnungszeit für Lehrer und Schüler sowie das Üben der verlangten Leistungen je zwei Unterrichtsstunden und der Hauptversuch drei Unterrichtsstunden beanspruchen. Es werden in etwa schwierigkeitsidentische Diktattexte aus den Vergleichsarbeiten von GLATZ verwendet, die Gedächtnisleistung und Konzentrationsfähigkeit prüfen, zum anderen Aufgabenreihen aus dem Zahlenrechnen, die Regelkenntnis der entsprechenden Rechenarten verlangen. Da im Rechenversuch schwierigkeits- und strukturidentische Parallelserien verwendet werden, müssen auftretende Wiederholungseffekte mit Hilfe von Kontrolluntersuchungen eliminiert werden.

Die Probesituationen sind so angelegt, daß sie vom Schüler zunehmende Anstrengung fordern, weil entsprechende Instruktionen seitens des Versuchsleiters immer bedeutungsvollere und weitreichendere Ziele ansprechen. Die Übungsarbeit ist dadurch gekennzeichnet, daß Leistungsfeststellungen weitgehend unterbleiben und eine relativ entspannte, leistungsdruckarme Atmosphäre erzeugt wird. Die Klassenarbeit fordert - im Rahmen der gewohnten Lehrerinstruktion und des üblichen Rituals - erhöhte Anstrengung und Aufmerksamkeit, weil Lehrerkontrolle, Zensur und Elternbenachrichtigung im allgemeinen mit dieser Arbeit verknüpft werden. Die Prüfungsarbeit zum Abschluß der Versuchsserie hat durch Gestaltung der Instruktion und des damit verbundenen Verhaltens des 'strengen Lehrers' den Charakter einer sehr gewichtigen, mitentscheidenden Abschlußarbeit, deren Einfluß auf die Zeugniszensur stark betont wird und die außerdem Aufstiegsmöglichkeiten in diesen Schuljahren anspricht (4. Schuljahr: Übergang zu weiterführenden Schulen; 9. Schuljahr: angestrebte Lehrstellen mit lukrativen Berufschancen, Weiterbildung auf Fachschulen).

Da die Auswirkungen von unterschiedlich anregenden Prüfungssituationen auf das Leistungsverhalten bestimmt werden sollen, müssen Zeitdruckerlebnisse wie auch Überlängen der Prüfarbeiten als Belastungsmomente konsequent ausgeschaltet werden. Daten, die Einstellungen zu einem Schulfach betreffen, die die Leistungsstärke eines Schülers näher bestimmen, werden erst während des oder nach dem Hauptversuch aufgenommen.

Die Auswertung der Befunde ergibt, daß die Schüler des 9. Schuljahres mit eindeutigeren Leistungsminderungen auf

die variierten Prüfungsbedingungen reagieren als Viertkläßler.. Alle Schülergruppen schneiden in der ungewohnten Situation der Prüfungsarbeit (PA) zum Teil wesentlich schlechter ab als in den Klassenarbeiten. Auch im Vergleich mit den Arbeitsergebnissen in der entspannten Übungsatmosphäre erleidet ein Teil der Schüler Leistungseinbußen, die allerdings nicht so eindeutig ausfallen wie die im Vergleich KA - PA. Die als sehr wichtig angekündigte, einmalige Prüfungsarbeit bringt den Schülern in der Regel also keine Leistungsvorteile, sondern eher mehr oder weniger deutliche Leistungseinbußen, wenn man ihre Ergebnisse mit den Daten der schulgewohnten Arbeitsproben vergleicht.

Die Schüler des 4. Schuljahres schneiden in Klassenarbeiten und Übungsarbeiten mit etwa identischen Fehlerraten ab. Die Instruktion der bedeutungsreichen Klassenarbeit wirkt sich also nicht leistungssteigernd aus, wenn man ihre Ergebnisse mit den Befunden der entspannten Übungssituation vergleicht. Im 9. Schuljahr dagegen darf man eine Überlegenheit der Klassenarbeit annehmen, denn im Rechtschreibversuch verbessern sich die Schüler recht eindeutig im Vergleich mit entsprechenden Übungsarbeitsergebnissen. Auch im Zahlenrechnen liegen Einzelbefunde (in den Einzelabschnitten) vor, die den leistungsfördernden Charakter der Klassenarbeitsinstruktion in diesem Schuljahr unterstützen. Insgesamt wirkt sich der Leistungsgewinn im Zahlenrechnen nicht so eindeutig aus wie im Diktatversuch. Einflüsse der Facheigenart und Schwierigkeit des Aufgabenmaterials können nicht beobachtet werden.

In der Analyse des Leistungsgeschehens erweist sich, daß die Einwirkung der experimentellen Bedingungen nicht schwerpunktmäßig lokalisiert werden kann, wenn der Gesamt-

versuch ins Auge gefaßt wird.

Der Einfluß der unterschiedlichen Prüfungssituation wirkt im Diktatversuch beider Schuljahre gleichmäßig auf Anfangs- und Endleistung.

Im Zahlenrechnen hingegen können deutliche Unterschiede zwischen den beiden Schuljahren in der Schwerpunktwirkung herausgearbeitet werden. Viertkläßler verschlechtern sich im Anfangsbereich (1. und 2. Einzelabschnitt) unter der Belastung bedeutungsreicher Arbeiten (ÜA - KA, KA - PA, ÜA - PA). In der Endleistung (3. und 4. Einzelabschnitt) hingegen läßt sich trotz leicht erhöhten Schwierigkeitsgrades der Rechenaufgaben keine leistungsfördernde Tendenz der Prüfungsbedingungen ausmachen.

Schüler des 9. Schuljahres (Zahlenrechnen) verschlechtern sich gerade in der Endleistung, wenn sie mit den ungewohnten Instruktionen der Prüfungsarbeit (PA) konfrontiert werden (ÜA - PA, KA -PA). Die Instruktionen der normalen schulischen Proben wirken sich nicht leistungsändernd aus. In der Anfangsleistung kann ein Einfluß der unterschiedlichen Prüfungsbedingungen nicht verzeichnet werden.

Die Qualität der Fachleistung scheint nur in der spezifischen Prüfungssituation der ungewohnten Prüfungsarbeit (PA) die Reaktionen unterschiedlich leistungsstarker Schüler auf die experimentellen Bedingungen zu differenzieren. Die Extremgruppen (leistungsstarke und leistungsschwache Schüler), besonders deutlich die leistungsschwachen Schüler, reagieren auf die Bedingungen der Prüfungsarbeit (PA) mit stärkeren Leistungsminderungen als die mittlere Leistungsgruppe, wenn man mit Arbeitsergebnissen der gewohnten schulischen Proben (ÜA,KA) vergleicht. Dieser Befund gilt nicht für den Diktatversuch des 4. Schuljahres. Einflüsse der Fachleistung können in diesem Zusammenhang nicht nachgewiesen werden.

Die Mädchen des 9. Schuljahres unterliegen im Zahlenrechnen stärker den Prüfungseinflüssen als Jungen. Dieser Befund wird durch ausgeprägtere Leistungsminderungen bestätigt, mit denen die Mädchen beider Schuljahre im Zahlenrechnen auf den Wechsel von gewohnten schulischen Leistungsproben (ÜA, KA) zur ungewohnten Prüfungsarbeit (PA) reagieren. Diktatergebnisse weisen diesen Trend nicht aus.

Hinsichtlich des Einflusses der Variablen 'Einstellung zum Schulfach' liegen nur wenige ausgeprägte Einzelergebnisse vor.

Ferner verdient noch die Tatsache festgehalten zu werden, daß die ungewohnte Prüfungsarbeit (PA) die Anzahl der Zeichensetzungsfehler nachhaltiger vermehrt als die Zahl der Wortfehler. Leistungsproben im üblichen Schulrahmen weisen diesen Unterschied nicht aus.

In leistungstheoretischen Überlegungen werden die Befunde im natürlichen Geschehenszusammenhang gesehen und gewichtet; Modellvorstellungen - allerdings mit noch unvollständigen empirischen Beweisketten - werden skizzenhaft angedeutet. Allgemein anerkannte Gesetzmäßigkeiten - wie das YERKES-DODSON-Modell - tragen dazu bei, erste Ansätze einer Theorie der Leistung in schulischen Prüfungssituationen deutlicher zu umreißen.

7. DISKUSSION DER ERGEBNISSE UND DES ANSATZES

7.1. Auseinandersetzung mit der Literatur

Alle Autoren, die sich experimentell mit dem Problem der seelischen Belastung in Prüfungssituationen auseinandersetzen, sind übereinstimmend der Meinung, daß der

moderne Mensch solche Situationen nicht vermeiden kann, da er von Geburt an den Einflüssen der an Leistung orientierten Umwelt ausgesetzt ist. Ebenso einig sind sich die Verfasser, daß gesteigerte Prüfungsbelastungen eine Erhöhung der Leistungsangst, der Aktivierung, der Empfindlichkeit nach sich zieht. MOELLER (1969) - psychoanalytisch orientiert - sieht einen positiven Effekt in der Prüfungsangst. Er glaubt an eine Motorfunktion der durch Prüfungen erzeugten Spannung für eine normale Entwicklung, wahrscheinlich auch mitangeregt von der Erkenntnis des YERKES-DODSON-Gesetzes, das bei geringem Aktivierungsniveau durch Prüfungsbelastung bis zu einem gewissen Grade Leistungsgewinne vorhersagt. Andere Aussagen von MOELLER (1969), wie auch ausgedehnte experimentelle Studien (HELM 1954, Arbeiten von J.G. SARASON und S.B. SARASON und ihrer Mitarbeiter) legen allerdings den Schluß nahe, daß das Leistungsbild durch Steigerung der Prüfungsbelastung über den gewohnten Rahmen hinaus eher von Leistungseinbußen bestimmt wird. Die vorliegenden Befunde sowie die Erhebung von SANDVOSS (1952) bestätigen diese Ergebnisse.

Die Mehrzahl der anglo-amerikanischen Autoren gewinnen ihre Befunde entweder mit schulunabhängigem Material, mit künstlich induziertem Stress oder an erwachsenen Menschen, deren Welt die Struktur der normalen schulischen Alltagssituation nicht widerspiegelt. Experimente mit Schülern der Haupt- und Grundschule mit schulischem Prüfungsmaterial im Schulalltag fehlen fast völlig. LEVITT (1971) vermutet, daß Aufwand, Labilität der Experimentiersituation (Kontrolle von intervenierenden Variablen), Schwerzugänglichkeit der Lehrer für psychologische Fragestellungen und Meßtechniken erfolgversprechende Experimente erschweren.

Die wenig umfangreiche Fachliteratur im deutschen Sprachbereich über das Problem des Prüfungsdrucks enthält neben sehr wenigen experimentellen Untersuchungen (SANDVOSS, HAECKER, HELM) vor allem Fallstudien sowie allgemein gehaltene ('all-round') Betrachtungen ohne experimentellen Hintergrund (ECKSTEIN 1958, TRAMER 1952 u.a. - siehe Kapitel 1.3.).

Innerhalb der für die vorliegende Fragestellung bedeutsamen experimentellen Ansätze gibt es nur zwei Untersuchungen, die das Problem des durch Lehrerinstruktion erhöhten Situationsdrucks angehen. SANDVOSS (1952) untersuchte das Leistungsverhalten von Schülern eines von ihm selbst geführten 2. Schuljahres (Versuchspersonen: 15). Er unterschied zwischen Übungs- und Klassenarbeit. Nach einer Einübungsphase sammelte er die Übungshefte ein. Er glaubte, damit ein Leistungsprodukt gewonnen zu haben, das in 'völliger Furchtlosigkeit', ohne Belastung zustande gekommen war. Am folgenden Tag wurde zur gleichen Stunde und unter sonst völlig gleichen Bedingungen die Probearbeit (Klassenarbeit) durchgeführt. Die Aufgaben (Additionen und Subtraktionen zweistelliger Zahlen mit gleichzeitigen Zehnerübergängen, veranschaulicht an der KUEHNELschen Hundertertafel) waren identisch im Schwierigkeitsgrad. Als Ergebnisse hält er fest:

a. "Der Druck der Probesituation nützt nur einer Minderheit, während die Mehrzahl der Kinder sich unter ihrem Einfluß wesentlich verschlechtert." (SANDVOSS 1952)

b. Die Lösungen einer Reihe fast gleichschwieriger Rechenaufgaben werden unter dem Druck der Probesituation gegen Ende immer fehlerhafter, während sie gegen Ende der in entspanntem Zustand erledigten Übungsarbeit besser werden.

Die in diesem Experiment auf ähnlichem Wege erzielten Ergebnisse bestätigen die erste Feststellung von SANDVOSS, wenn man annimmt, daß zu Beginn des 2. Schuljahres auch Klassenarbeiten noch zu den Prüfungssituationen gehören, die dem Kind relativ unbekannt und ungewohnt sind und deshalb auch gesteigerte Leistungsangst auslösen, die etwa vergleichbar mit der schulunüblichen Prüfungsarbeit im 4. und 9. Schuljahr Leistungseinbußen mit sich bringt, wenn man die Ergebnisse der Übungssituation als Vergleich heranzieht. Für die Bestätigung des zweiten Ergebnisses können eindeutige Befunde nicht genannt werden.

In einem anderen Beitrag befaßt sich HAECKER (1964) ebenfalls mit der unterschiedlichen Auswirkung von Übungs- und Klassenarbeitssituationen auf das Leistungsverhalten von Volksschülern. HAECKER untersuchte 306 Schüler einer achtklassigen Volksschule (pro Schuljahr zwei Klassen). Unter Einbeziehung einer Übungsphase ließen die jeweiligen Fachlehrer - mit HAECKER als stillem Beobachter - unter sonst gleichen Bedingungen eine Übungsarbeit und eine Klassenarbeit schreiben. Außerdem wollte HAECKER durch zusätzliche Befragung das kindliche Erleben der Prüfungssituation aufhellen. Als Prüfungsmaterial verwendete er Aufgaben des reinen Zahlenrechnens mit durchgängig identischem Schwierigkeitsgrad (bezogen auf die Aufgabenserie eines Jahrganges).
Als Ergebnis hält HAECKER fest:
"Die Kinder des 1. und 2. Schuljahres wiesen eine sehr signifikante Leistungsverbesserung auf. Im 3. bis 7. Schuljahr zeigte sich die Tendenz zur Leistungsverschlechterung, die im 3. und 7. Schuljahr sehr signifikant war".
(HAECKER 1964)

Anhand der Ergebnisse der vorliegenden Felduntersuchung können die von HAECKER registrierten Leistungsunterschiede zwischen Übungsarbeit und schulgewohnter Klassenarbeit

nicht als bestätigt angesehen werden.

Die Kritik an den Unterrichtsversuchen von SANDVOSS und HAECKER entzündet sich einmal an der geringen Anzahl der Versuchspersonen pro Schuljahr (SANDVOSS 15 Vpn, HAECKER 22 Vpn), zum anderen an der Vernachlässigung von Wiederholungseffekten bei gleichschwierigen Aufgabenserien im Zahlenrechnen (Kontrolluntersuchungen bestätigen solche Effekte). Andeutungen von HAECKER ("die Anzahl der innerhalb einer Unterrichtsstunde von den meisten Kindern zu bewältigenden Aufgaben"), das Ausmaß der in 45 Minuten zu bewältigenden Aufgabenmenge (z.B.: 4. Schuljahr: 20 Aufgaben wie: 25546 : 7) und fehlende Hinweise auf Ausschalten von Zeitbeschränkungen am Ende der Unterrichtsstunde lassen den Verdacht als begründet erscheinen, daß Phänomene des relativ milden Zeitdrucks die Qualität der schulischen Leistung in seiner Untersuchung mitbeeinflussen. Zudem soll auf die möglichen, auch durch Verhaltensabsprachen nicht auszuschaltenden Einflüsse der Lehrerpersönlichkeiten, die HAECKERs Experiment durchführten, hingewiesen werden. SARASON (1954) und SCHAFER (1954) bestätigen, daß soziale Beziehungen zwischen Versuchsleiter und Versuchsperson einen entscheidenden Einfluß auf das Leistungsverhalten der Probanden haben können. Ähnliche Befunde erwähnen LEWIN, LIPITT, WHITE (1939), COGAN (1956), TAUSCH, KOEHLER, FITTKAU (1966).

Beiträge aus dem umfangreichen Schrifttum der Stress-Forschung sollen im folgenden nur soweit referiert werden, wie sie eine Bedeutung für die im Rahmen der vorliegenden Fragestellung erlangen.

Im Experiment werden die Effekte der unterschiedlich belastenden Prüfungssituationen auf die Schulleistung möglichst schulnah - unter Verwendung von unterrichtsspezifischem Material in der schulischen Alltagssituation - analysiert. LEVITT und HOLZKAMP betonen die Fruchtbarkeit solcher Ansätze. Nach LEVITT (1971) und HOLZKAMP (1972) kommt es angesichts der geringen Validität von Experimenten mit isolierten, reduzierten und labilisierten Bedingungen, deren Effekte in der praktischen Unterrichtssituation kaum eine berücksichtigungswerte Größenordnung besitzen und auch kaum auf eine unreduzierte, komplexe Alltagssituation übertragen werden können, vor allem darauf an, "Strukturen typischer Realitätsformen zu erfassen, mit denen der Praktiker konfrontiert ist" (HOLZKAMP 1972) und sie, wenn eben möglich, in einer realitätsnahen Umwelt zu erkunden. Es ist also für das Wissen um die Auswirkung schulischer Prüfungsbelastungen enorm wichtig, daß sie auch in der realen Umgebung, in der sie auftreten, untersucht werden, obwohl, wie auch in diesem Experiment ersichtlich, sich dabei "methodologische Probleme der Standardisierung von Verfahrensweisen und Kontrolle der äußeren Faktoren" (LEVITT 1971) gegenüber schulfernen Experimenten erheblich vergrößern.

Untersuchungen von GRINKER u.a. (1957) sowie von WALKER und SPENCE (1964) über das Erleben von künstlich erzeugtem Stress in Laboratoriumsexperimenten lassen erhebliche Zweifel an der Gültigkeit der Experimente mit solchen Stresstechniken aufkommen und bestätigen die Einwände HOLZKAMPs (1972) gegenüber einem Bereich der modernen psychologischen Forschung, deren Bemühen in erster Linie der "Verfeinerung der Design- und Meßtechniken und der Interferenz-Statistik gilt, und andere Kriterien, wie äußere Relevanz, vernachlässigt" (HOLZKAMP 1972).

Die Deutung der Befunde geht vom Gedanken aus, daß die Schüler im Laufe einer langen Entwicklung die Grundstrukturen des Leistungssystems ihrer Umwelt kennenlernen und gegenüber Leistungssituationen sehr sensibel reagieren. Der Ursprung einer solchen Sensiblisierung muß in der Vergangenheit des Kindes gesucht werden.

Diese Erkenntnis wird bestätigt durch die Längsschnittuntersuchungen von HILL, SARASON (1966), die hervorheben, daß die Empfindlichkeit gegenüber schulischen Leistungsproben nicht zuletzt eine Funktion der schulischen und häuslichen Eindrücke ist, die das Kind im Laufe seiner Entwicklung aufgenommen hat.

S.B. SARASON u.a. (1960) sehen die Ursache der hohen Empfindlichkeit Prüfungssituationen gegenüber - wohl von FREUDschen Konzeptionen beeinflußt - in einer stark gestörten Eltern-Kind-Beziehung, die während prüfungsähnlicher Situationen im Unbewußten neu belebt wird.

Nach den Literaturbefunden spricht also viel für die Annahme, daß früher erlebte und erlernte Verhaltensweisen gegenüber Leistungsansprüchen, -erwartungen und -forderungen der Erzieher (Elternhaus, Kindergarten, Schule) durch situative Momente in der aktuellen schulischen Prüfungssituation (Bekanntheitsgrad der Situation und der Aufgaben, Instruktionen, Verhalten des Lehrers, Bedeutung der Prüfung - studienbegleitend oder alles entscheidend - sowie die in dieser Untersuchung nicht präziser nachgeprüften Variablen der Schwierigkeit, der Lösbarkeit und der Komplexität des Prüfmaterials) neu belebt werden und eine sachimmanente Zuwendung als Voraussetzung der erfolgreichen Schulleistung nur selten erleichtern (9. Schuljahr, Diktat, ÜA-KA), oft aber auch behindern (die fehlererhöhende Wirkung der ungewohnten Prüfungsarbeit)

Die vorliegenden Befunde einschließlich der von SANDVOSS beigesteuerten Ergebnisse rechtfertigen die Annahme, daß das durch die Länge des Schulbesuchs geprägte Erleben von situativen Faktoren der schulischen Prüfungssituationen das Leistungsverhalten der Schüler wesentlich mitbestimmt. So ist aus Experimenten - allerdings mit Studenten - bekannt, daß der Vertrautheitsgrad einer Prüfungssituation mit ausschlaggebend für die Angst und die Höhe der erreichten Leistung ist (CASTAN, GRENIER 1967, ZIMBARDO, BARNARD, BERKOWITZ 1963). Weiterhin untermauern SARNOF u.a. (1958) zahlreiche Befunde in Experimenten mit Studenten, wenn sie hervorheben, daß die Form studienbegleitender kleiner Prüfungsarbeiten, wie sie die Klassenarbeiten darstellen, zumindest nach längerer Eingewöhnungszeit angstreduzierend und leistungsfördernd wirkt gegenüber einem System, das nur alles entscheidende Abschlußexamen kennt.

Zwei Konzepte stehen im Mittelpunkt der theoretischen Erörterungen von Leistung und Angst in Prüfungssituationen. Ausgangsbasis aller Überlegungen ist das lerntheoretische Konzept von HULL (1943), das zugleich die Grundlage für das Modell von SPENCE (1958) und TAYLOR (1956) bildet. Dieses Konzept geht von der 'habit-Bildung' als einer Verknüpfung von Reiz und Reaktion im THORNDIKEschen Sinne aus, erkennt aber auch intervenierende Variable zwischen Reiz und Reaktion an, wie Stärke der Verknüpfung, Bedürfnislage, primäre Motivation. Im Leistungsprozeß haben solche Verbindungen Erfolg, die zu einer Problemlösung führen. SPENCE behauptet, daß es im Laufe der Leistungsprozesse zur Ausbildung einer 'habit-Hierarchie' kommt, wobei die hoch in der 'habit-Hierarchie' stehenden richtigen Antworten die geforderten Leistungen erleichtern, besonders dann, wenn das Angstpotential des Probanden, wie SPENCE vermutet, jedem 'habit' in dieser Hierarchie

proportional zu seiner ursprünglichen Stärke Energie zuführt. Es wird folgende multiplikative Beziehung postuliert: aktiviertes Angstpotential im Organismus als Antriebsenergie mal 'habit-Stärke' (Reaktion E = D mal H). In sehr einfachen Lernaufgaben, wie zahlreiche Lidschlußexperimente beweisen (J.A. TAYLOR 1951, 1953; SPENCE und FARBER 1954), liefern tatsächlich Hochängstliche bessere Leistungen bei der Vermeidung eines Reizes (HILGARD 1951: meist angsterregender Störreiz) als weniger ängstliche Personen. Menschliche Leistungsprozesse, in denen kognitive Fähigkeiten angesprochen werden, sind allerdings komplexe Bedingungsgefüge mit weit mehr als nur einem 'habit'. Da immer viele konkurrierende 'habits' in solchen komplexen Leistungsprozessen mitwirken, stehen Handlungstendenzen in diesem Bereich meist relativ niedrig in der 'habit-Hierarchie' und begünstigen bei Aktivierung des Angstpotentials Fehllösungen.

Insgesamt kann man sagen, daß dieses auch als Iowa-Theorie bekannte Modell von SPENCE und TAYLOR die Erregung von Angst als Funktion eines menschlichen Wesenszuges versteht. Situative Momente haben nach SPENCE und TAYLOR hinsichtlich der Aktivierung des Angstpotentials nur sekundäre Bedeutung.

MANDLER und SARASON (1952) glauben anhand experimenteller Studien eine andere Theorie, die sogenannte Yale-Theorie, vertreten zu können, die kausale Zusammenhänge zwischen "test-anxiety" und kognitiven Prozessen annimmt und das "drive-Modell" von SPENCE und TAYLOR angesichts der unüberwindlichen Schwierigkeit, 'habit-Systeme' für Problemaufgaben zu analysieren, zurückweist. MANDLER und SARASON gehen von der auch diesem Experiment zugrunde gelegten Erfahrung aus, daß Leistung und damit zusammen-

hängende Verhaltenskonsequenzen im Leben des in einer leistungsorientierten Umwelt aufgewachsenen Menschen eine entscheidende Rolle spielen. Sie ersetzen den Begriff Leistungsangst durch Testangst ("test-anxiety"), um ihn einzugrenzen und zugänglicher zu machen (vgl. die zahlreichen, in den USA entwickelten Angstfragebogen). In dieser Untersuchung werden die Begriffe Leistungsangst, Prüfungsangst und Testangst synonym gebraucht, weil sie sich eindeutig auf Prüfungssituationen beziehen, wie sie in der Schulpraxis und Studium üblich sind.

Nach MANDLER und SARASON hängt die Angstentwicklung mehr von den situativen Faktoren in Prüfungssituationen wie Instruktion, Bekanntheitsgrad der Situation und der Aufgabe, Bedeutung der Probesituation, Verhalten des Lehrers und Struktur des Aufgabenmaterials ab. MANDLER und SARASON unterscheiden zwei Antriebe in Leistungssituationen: einmal die "task-drives", die auf Erledigung der Aufgaben drängen, und die "anxiety-drives", deren eine Komponente aufgabenrelevante Reaktionen enthält, die zur Lösung der Bedrückung führen sollen, also zur Problemlösung beitragen, deren andere Komponente aber aufgabenirrelevant, d.h. leistungshemmend wirkt. "Anxiety-drives" beeinflussen die Leistung negativ, wenn Gefühle der Unzulänglichkeit, der Hilflosigkeit, Fluchttendenzen, Angst vor Versagen und somatische Beschwerden in Prüfungssituationen überhand nehmen. Belastungen in Prüfungssituationen führen immer zu Interferenzen von aufgabenrelevanten und -irrelevanten Reaktionstendenzen.

Eine wichtige Grundlage der Yale-Theorie ist noch anzufügen. MANDLER und SARASON glaubten, daß die "anxiety-drives" nur in einer von zwei entgegengesetzten Richtungen wirken. ALPERT und HABER (1960) fanden heraus, daß beide Richtungen der Angstreaktionen im Individuum nebeneinander bestehen und interferieren können.

Die vorliegende Untersuchung lehnt sich in ihrem Ansatz mehr an die von ALPERT und HABER ergänzte Yale-Theorie von der situationsbedingten Auslösung von Leistungsangst in Belastungssituationen an.

Die kaum voneinander abweichenden Fehlerraten der in entspannter Atmosphäre geschriebenen Übungsarbeit und der gewohnten, aber doch bedeutungsvollen Klassenarbeit (Ausnahme: 9. Schuljahr, Diktat) können so erklärt werden, daß mit zunehmender Belastung neben den intensiven, aufgabenbezogenen Reizen, die rasche Erledigung der Aufgaben fordern ("task-drives"), die aufgabenirrelevanten Reaktionen der "anxiety-drives" die Leistung in gewohnten, durchschaubaren Situationen nicht negativ beeinflussen, da sich die beiden Komponenten der "anxiety-drives" in ihrer Wirkung in etwa aufheben. Das Leistungsniveau bleibt trotz erhöhter Belastung gleich. Der Befund im Diktatversuch des 9. Schuljahres, nach dem Schüler in der Klassenarbeit besser abschneiden als in der Übungsarbeit, läßt nur den Schluß zu, daß die in der Übungssituation wirksamen "task-drives" durch aufgabenrelevante Reaktionen der "anxiety-drives" in ihrer Wirkung (leistungsfördernd) verstärkt werden. Leistungshemmende Einflüsse der interferierenden "anxiety-drives" setzen sich in der gewohnten schulischen Prüfungssituation der Klassenarbeit nicht eindeutig durch.

Nach der Annahme des YERKES-DODSON-Gesetzes ergibt das Zusammenwirken von "task-drives" und aufgabenrelevanten "anxiety-drives" bei etwa mittlerer Belastung ein Leistungsoptimum; mit zunehmender Belastung setzen sich die leistungshemmenden Effekte der "anxiety-drives" gegenüber den aufgabenrelevanten Reaktionen durch. Diese Vermutung wird bestätigt durch den statistisch gesicherten, nachhaltigen Leistungsabfall in der ungewohnten, hinsichtlich

der Konsequenzen nicht durchschaubaren, sehr bedeutungsvollen Prüfungsarbeit.

Im Sinne der Iowa-Theorie müßten, da nur relativ einfache Leistungsprozesse analysiert werden, kontinuierliche Leistungssteigerungen zu beobachten sein. Die Auslöseleichtigkeit E einer Reaktion erhöht sich nämlich unter der Voraussetzung, daß in der multiplikativen Beziehung E = D (drive) mal H (habit-strength) der Faktor H (identisches Prüfmaterial) gleichbleibt und der Faktor D (Angstpotential, Angstgrad, drive) durch den Einfluß der experimentellen Bedingungen größer wird (vgl. auch Tierexperimente von MILLER 1951, der durch künstlich induzierte Furcht erhöhte Leistungen erzielte).

Die Iowa-Theorie kann die Auswirkung der Variablen Leistungsangst, Belastung nur in relativ einfachen menschlichen Leistungsprozessen (Konditionieren) hinreichend erklären. Liegen solche Leistungen, wie Rechtschreiben und Zahlenrechnen vor, gelangt sie nur unter Einbeziehung zahlreicher, z.T. nicht verifizierbarer Zusatzannahmen (Problem der habit-Hierarchie) zu Erklärungen. Die von ALPERT und HABER ergänzte Yale-Theorie der situationsbedingten Angstauslösung kann unter Einbeziehung der Gesetzmäßigkeiten des YERKES-DODSON-Modells auch die Befunde der vorliegenden Untersuchung einleuchtend erklären.

MANDLER und SARASON fordern die eingehende Untersuchung spezifischer Stress-Situationen, wie es im vorliegenden Experiment geschieht, um dem Problem der Angst in Leistungssituationen gerecht zu werden.

7.2. Pädagogische Erörterungen: Reduzierung von Angst- und Belastungszuständen im Schulalltag

Die experimentellen Befunde legen die Vermutung nahe, daß die Situation der Klassenarbeit, eine vertraute schulische Belastungssituation, nur selten - bei unterrichtserfahrenen Entlaßschülern der Hauptschule - zu Leistungsgewinnen beiträgt (nicht durchgängig), bei jüngeren Schülern (4. Schuljahr) das Leistungsniveau nicht hebt und bei Schulneulingen (so darf mit SANDVOSS 1952 auch noch zu Beginn des 2. Schuljahres angenommen werden) das Leistungsbild gegenüber Ergebnissen in leistungsdruckfreien Situationen wohl negativ beeinflußt. Ungewohnte, harte, strenge, einmalig wichtige Prüfungsbelastungen wirken sich in diesem Sinne durchgängig leistungsmindernd aus. Diese Erkenntnis kann, da sie mit schulabhängigem Prüfmaterial im Schulalltag gesichert wurde, anregen, schulpraktische Folgerungen zu überdenken, die das Schülerdasein vom Druck, von der Angst und Sorge in schulischen Leistungskontrollen befreien helfen und dazu beitragen, Leistungsfähigkeit und Leistungsfreude zu erhalten. Die Möglichkeit, seelische Belastungen im Schulalltag überhaupt zu vermeiden, entfällt, weil, wie in der Einführung betont wird, die Schule und die in ihr tätigen Lehrer von den Bedürfnissen und Forderungen der Umwelt geleitet werden. Außerdem sei an die außerschulischen Erziehungsträger erinnert, deren Mitglieder ebenfalls Repräsentanten der leistungsorientierten Umwelt sind. Das häufig lautgewordene Verlangen, in der modernen Welt Angst- und Belastungszustände gänzlich abzubauen, läuft auf eine radikale Abkehr vom Leistungs- und Erfolgsprinzip hinaus, ein Vorgang, der fundamentale Grundlagen des Gesellschaftssystems in Frage stellen würde und angesichts

entsprechend unüberwindbarer Widerstände unrealisierbar ist.

Es kann also derzeit nicht darum gehen, Leistungsproben grundsätzlich abzuschaffen, sondern es müssen Wege und Möglichkeiten der Angst- und Belastungsreduzierung gesucht werden, die trotz aller Hindernisse praktikabel und wirksam sind. Der vorher zitierte Schulpraktiker SANDVOSS (1956) verweist auf einen weiteren, von ihm durchgeführten Unterrichtsversuch in einer Klasse für Lernbehinderte, der einige wertvolle Anregungen für die Gestaltung eines belastungsreduzierenden Unterrichts anbietet: Vermeiden allzu hoher Leistungserwartungen seitens des Lehrers, keine Klassenarbeiten im herkömmlichen Sinne, sondern etwa in Form von Übungsarbeiten in völlig entspannter, leistungsdruckfreier Situation, individuelle Förderung von rechenschwachen Kindern (also keine Bildung von Leistungsgruppen), Ausschaltung der Kontrolle der Hausaufgaben durch die Eltern, Lob und Ermutigung für jede Schülerleistung, Herunterspielen von Schülerversagen (es macht gar nichts, wenn Du mal eine Sache nicht kannst). Nach konsequenter Durchführung dieser Prinzipien bemerkte SANDVOSS etwa ein halbes Jahr später einen Leistungsanstieg und registrierte mit 'Erstaunen', daß gerade krasse Rechenversager nun relativ sicher und mit viel mehr Begeisterung als vorher rechneten.

Dieser Katalog konkreter Hinweise ist zu ergänzen mit Blick auf die Eigentümlichkeiten des heutigen Grund- und Hauptschulsystems: Änderung des derzeit praktizierten Systems der Leistungsgruppendifferenzierung in der Hauptschule (Vermeiden gettoähnlicher Zustände für die unterprivilegierten B- und C-Gruppen), Schwerpunktverlagerung vom Frontalunterricht auf Gruppenarbeit mit Schülern aller Begabungsstufen und individuellen Förderungsmöglichkeiten (Experimente von SCHACHTER 1959 und WRIGHTSMAN 1960 be-

stätigen Möglichkeit der Angstreduzierung in Leistungssituationen durch echte Mitgliedschaft in einer Gruppe), intensive Betreuung der Schüler durch einen Lehrer.

Alle diese Maßnahmen laufen auf ein auch von SANDVOSS durchgängig praktiziertes Erziehungsprinzip hinaus: "Entangsten" (SANDVOSS 1956) der schulischen Alltagssituation durch Ermutigung, durch Steigerung des Selbstvertrauens, des Selbstwertgefühls und der Selbstsicherheit (vgl. auch ADLER, KUENKEL 1928, PLATTNER 1968). Allerdings verliert dieses hoffnungsvolle Konzept an Wirkung, wenn nicht entscheidende vor- und außerschulische Erziehungsträger (vor allem die Eltern der Kinder) über Ursachen und Auswirkungen von Angstreaktionen, die von den elterlichen Erziehungsmaßnahmen ausgehen, unterrichtet werden und zugleich in Beratungs- und Trainingsstunden die konkrete Wirkung eines Erziehungskonzeptes selbst erleben.

Einige Experimente, die sich mit der Angstreduzierung in Belastungssituationen auseinandersetzen, liefern interessante Erkenntnisse, die auch für die tägliche Unterrichtsgestaltung als wertvolle Hilfen angesehen werden können und helfen, leistungshemmende Angstreaktionen auszuschalten oder zumindest abzubauen.

So versuchten PAUL und ERIKSEN (1964) Prüfungsangst mit Hilfe spezieller Instruktionen zu reduzieren, was ihnen auch gelang. Der Lehrer kann also nicht nur, wie in diesem Experiment praktiziert, Belastung und Angst durch Instruktionen im Schüler verstärken, sondern auch verkrampfte, belastete Schüler mit freundlichen, ermutigenden Worten in eine Stimmung versetzen, die dazu beiträgt, daß der geprüfte Schüler eine Leistung erbringt, die seiner tatsächlichen Leistungskapazität in etwa entspricht.

PAUL und ERIKSEN bestätigen mit einem weiteren Befund SANDVOSS (1956), wenn sie betonen, daß es darauf ankommt, den Schülern nicht immer die Wichtigkeit eines guten Abschneidens oder - noch schlimmer - die Folgen eines Versagens in Form öffentlicher Bloßstellung vor Kameraden und Eltern vor Augen zu stellen.

Nach METZGER (1957) ist die Einführung neuer unbekannt Schwierigkeiten in schulischen Leistungsproben ein bevorzugtes Druckmittel des 'strengen Lehrers'. Experimentelle Untersuchungen von ZIMBARDO, BARNARD, BERKOWI (1963) bestätigen die vermuteten angsterhöhenden und leistungsmindernden Wirkungen einer solchen Handlungsweise. Sie fanden erhöhte Angstreaktionen in sogenannten offenen Situationen, in denen es den Probanden nicht klar war, was erwartet wurde und welche Aufgaben zu lösen waren.

GIFFORD und MARSTON (1966), CASTAN und GRENIER (1967) stellen übereinstimmend fest, daß Probanden weniger Angst entwickeln, wenn sie wissen, was sie erwartet, wenn Übungsphasen oder orientierende Sitzungen der Ernstsituation vorausgehen. Die Leistungssituation muß in jeder Hinsicht durchschaubar sein, ohne daß der Lehrer als 'Geheimniskrämer' fungiert.

BIGGS (1962) weist auf den Tatbestand hin, daß "traditionelle Methoden", also Vorgehensweisen, die mechanisches, regelhaftes, schematisches Lernen fördern, eher Schulangst hervorrufen als eine Unterrichtsgestaltung, die auf Einsicht, Überblick, vorurteilslose Betrachtung der Ausgangslage Wert legt.

SARASON u.a. (1960) berichten, daß beliebte Lehrer gerade dann Angst erzeugen, wenn sie die Akzeptierung eines Schülers von dessen Leistung abhängig machen. Sie betonen wie SANDVOSS u.a. den Wert der Ermutigung im Falle des Schülerversagens. Der Lehrer soll aktiv eingreifen und nicht versuchen, die Leistungsschwäche durch Abwenden vor den anderen zu verbergen, denn nach SARASON u.a. wird eine solche Handlungsweise vom betroffenen Schüler gewöhnlich als Bestrafung angesehen.

Als letztes Experiment, das Erziehern Hinweise für konkrete angstreduzierende Handlungen anbietet, soll auf das bekannte Experiment von LEWIN, LIPPITT und WHITE (1939) hingewiesen werden, das die Wirkung autoritärer und demokratischer Führungsstile auf das Verhalten von Schülern analysiert. Neben einer Leistungsüberlegenheit fällt vor allem der Abbau von Verzagtheit und die Förderung der "Antriebsquellen" (METZGER 1971) in der demokratisch geführten Gruppe auf, in der zwischen Versuchsleiter und Schülern ein partnerschaftliches Verhältnis im Dienste des zu bewältigenden Problems herrschte. Alle methodisch-didaktischen Hinweise zielen auf ein bestimmtes Lehrer - Schüler - Verhältnis ab. Leistungsproben in der Schule können nicht vermieden, aber erleichtert, weniger belastend gestaltet werden. Der Lehrer kommt seiner Erziehungsverpflichtung in dieser Hinsicht am ehesten nach, wenn er sich als partnerschaftlicher Helfer des Schülers betrachtet und mit ihm Probleme meistert.

8. AUFGABEN FÜR DIE WEITERE FORSCHUNG

Mit der vorliegenden Untersuchung ist der Nachweis erbracht, daß das Problem der Leistung in schulischen Prüfungssituationen ein fruchtbarer Ansatz der Unterrichtsforschung ist und Ergebnisse ergänzt, die einmal in Laboratoriumsexperimenten mit künstlich induziertem Stress und zumeist schulunabhängigem Material gefunden wurden, die zum anderen aus Untersuchungen mit Studenten in der extrem belastenden Abschlußprüfung eines Studienabschnittes resultierten.

Eine Erweiterung der Untersuchung auf die übrigen Klassen der Grund- und Hauptschule könnte vermutlich das skizzenhaft angedeutete Modell der Leistung in schulischen Prüfungssituationen differenzieren. Zudem kann die Frage beantwortet werden, wie die zeitliche Nähe von Zeugnissen, und besonders Abschlußzeugnissen, die schulische Leistung und das Erleben von Prüfungssituationen hinsichtlich einer gesteigerten Sensibilisierung verändert. Die bezüglich ihres Effektes umstrittene Leistungsdifferenzierung in einigen Fächern des Hauptschulunterrichts (A-, B-, C-Gruppierung) könnte auf Grund empirischer Ergebnisse zum Thema: Leistung und Angst in schulischen Prüfungssituationen überdacht und evtl. im Sinne einer besonders für Leistungsschwache angstreduzierenden Gestaltung von Gruppenarbeit (z.B. innere Differenzierung in Klassenverbänden) revidiert werden.

Die Ausdehnung des Ansatzes auf andere Schularten - wie Gymnasium und Sonderschule - könnte interessante Zusammenhänge von Struktur der Prüfungssituation und Eigenheit des betreffenden Schulsystems aufdecken. So sind etwa Prüfungsvorgänge im stark leistungsorientierten Unter-

richt der Höheren Schule zu studieren. Auf der anderen Seite dürfte das Verhalten lerngestörter Kinder in Leistungssituationen wichtige Hinweise auf Möglichkeiten der Verhütung von Prüfungsdruck liefern.

Es wäre ebenfalls reizvoll, im Rahmen laufender Gesamtschulversuche zu prüfen, wie die schon vor Jahren angestrebte und zum Teil vollzogene Auflösung der starren Bildungsgänge (Durchlässigkeit der Schulsysteme, Auflösung der Jahrgangsklassen, Einrichtung von Neigungsgruppen, Aufbau musischer Zweige des Gymnasiums) die Angstentwicklung, das Wesen, den Ablauf und das Ergebnis von schulischen Prüfungsvorgängen beeinträchtigt.

Vermutete Beziehungen zwischen Art der Probesituation und Struktur des Aufgabenmaterials können dahingehend analysiert werden, daß nicht nur relativ einfache Leistungsprozesse (wie in dieser Untersuchung Zahlenrechnen, Rechtschreiben) geprüft werden, sondern auch Problemaufgaben (Textaufgaben, Alltagsprobleme), deren Lösung Vorausplanung, Begriffsbildung, unkonventionelles, originelles, vom Schema abweichendes, vorurteilsloses Vorgehen erfordert. Das Prüfmaterial kann weiterhin differenziert werden nach Schwierigkeit, Komplexität und Lösbarkeit, wie es DUNN (1968) recht anschaulich vorgeführt hat. Hierbei würde eine Verhaltensbeobachtung zusätzlichen Aufschluß über die Struktur psychischer Prozesse in Prüfungssituationen geben können. Andere Indikatoren von Belastung und Angst, die Reaktionen des autonomen Nervensystems (Blutdruck, Pulsfrequenz, Atmung u.a.), entfallen als Meßgrundlagen, da solche individuellen Messungen während des Prüfungsvorganges die situativen Faktoren des alltäglichen Bedingungs-

zusammenhanges in der Schule verändern und die "Intimität der pädagogischen Situation" (WINNEFELD 1957) zerstören würden.

Im Sinne der Angsttheorie von MANDLER und SARASON (1952) kommt es darauf an, unterschiedlich strukturierte Stress-Situationen im Schulalltag ohne Verletzung der natürlichen Geschehenszusammenhänge in ihren Effekten auf das Leistungsverhalten der Schüler zu vergleichen. So ist bisher ungeklärt, ob die Bedingungen der schriftlichen schulischen Prüfung mehr Fehler und Angst hervorrufen als bestimmt strukturierte mündliche Prüfungen. Fragen bleiben offen wie: Welche Struktur muß eine mündliche Prüfung aufweisen, damit die Mehrheit der Schüler ein Leistungsoptimum erreicht? Wie wirken sich schulische Prüfungssituationen mit differenzierter Benotung in Form von Gutachten aus, die individuelle Schwierigkeiten des einzelnen Schülers auf dem Weg zur Lösung mitbeachten und auswerten, gegenüber Bewertungen mit dem gewohnten Sechser-Noten-System oder Alternativurteilen (bestanden, nicht bestanden)?

Der Einfluß relevanter Persönlichkeitszüge wie Ängstlichkeit, Leistungsmotivation, Schulbegabung und neurotische Neigung auf die Leistungen in unterschiedlich belastenden schulischen Prüfungssituationen ist experimentell nicht nachgeprüft. Unterrichtsversuche werden diese Zusammenhänge klären können.

Beachtung verdienen auch andere unterrichtsspezifische Variable wie Lehrerverhalten, Führungsstil, Unterrichtsmethodik, die den Geschehenszusammenhang im schulischen Alltag wesentlich mitbeeinflussen. Großzügige Unterrichts-

versuche, in denen die gruppenbezogene Zusammenarbeit der Schüler den üblichen Frontalunterricht ablöst, können die Nützlichkeit von Gruppenprüfungen im Vergleich mit den Effekten des heute praktizierten Einzelprüfsystems bestimmen.

Ein noch ungeklärtes, aber äußerst wichtiges Problem ist die Wirkung von unterschiedlich strukturierten Prüfungssituationen auf Schüler, die in der gruppenbezogenen Kooperation erfahren sind und Schüler, die im üblichen Frontalunterricht ausgebildet sind. WITTOCH (zit. nach METZGER 1971) entdeckte nach ihren Unterrichtsversuchen Verhaltensunterschiede in abschließenden Prüfungen: Es zeigte sich, "daß sich die Mitglieder der im Frontalunterricht geführten Gruppen bei der Lösung von Textaufgaben, die Umstrukturierungen erforderten, zum größten Teil nach Hilfe umschauten; die Mitglieder der Erarbeitungsgruppe, die vorher gemeinsam am Werk gewesen waren, legten Wert darauf zu versuchen, ob sie es allein könnten". (S. 185 - 187)

Diese Anzeichen deuten darauf hin, daß eine gruppenbezogene Unterrichtsgestaltung, in der der Lehrer seine übliche Gängelei aufgibt und das 'Denken und das Entdecken von Lösungsstrukturen seinen Schülern überläßt', die Leistungsangst in Prüfungssituationen zu reduzieren hilft. Der Schüler hat gelernt, Probleme auf dem Weg zu einem mehr oder weniger attraktiven Ziel als Hindernisse zu erleben, die er selbst überwinden kann. Die Prüfungssituation enthält auch Probleme, die problembewußt und sachzentriert gelöst werden können, ohne daß der Schüler - er kann in diesen Leistungssituationen nicht darauf hoffen, daß der Lehrer ihn über die Hindernisse trägt - verzagt, mutlos und verängstigt wird.

Die fundamentale Bedeutung der Lehrerinstruktion für Stress-Induzierung und -Reduzierung während schulischer Prüfungssituationen sollte dazu anregen, Trainingsprogramme auszuarbeiten, die die Mitverantwortung des Lehrers für das Leistungsverhalten der Schüler herausstellen und ihm Wege aufzeigen, wie er überhöhten, leistungsmindernden Prüfungsdruck verhindert (vgl. Kap. 7.2.)

Der vorliegende Forschungsbereich: Angst und Leistung in schulischen Prüfungssituationen ist in empirischen Forschungsvorhaben nur selten beleuchtet worden. Es bieten sich mehrere Forschungsansätze an (vgl. Kap. 1.2.), von denen die vorliegende Arbeit eine Möglichkeit der Verwirklichung ist.

LITERATUR

ALPERT, R., HABER, R.N. (1960): Anxiety in academic achievement situations. J. abn. soc. Psychol. 66, 207-215

ATKINSON, J.W., LITWIN, G.H. (1960): Achievement motive and test anxiety conceived as a motive to approach success and motive to avoid failure. J. abn. soc. Psychol. 60, 52-63

BARTMANN, TH. (1963): Der Einfluß von Zeitdruck auf die Leistung und das Denkverhalten von Volksschülern. Psychol. Forschung 27, 1-61

BECK, W. (1953): Grundzüge der Sozialpsychologie. München

BIGGS, J.B. (1962): Anxiety, motivation and Primary school mathematics. The National Foundation for Educational Research in England and Wales, Occasional Publication No.7

BROADHURST, P.L. (1959): The Yerkes-Dodson- law revived. Acta psychol. (Amst.) 16, 321-338

BROEN, W.E. (1959): Anxiety, intelligence, achievement. Psychol. Repts. 5, 701-704

CASTAN, P., GRENIER, M. (1967): Role de l'accueil dans l'examen psychotechnique de candidates a une formation professionelle. Cahiers de Psychologie 10, 151-160

COGAN, M.L. (1958): The Behavior of Teachers and The Productive Behavior of Their Pupils. J. exp. Educ. 27, 89-124

CORRELL, W. (1962): Erziehung zur Angsttoleranz. Lebendige Schule 17, 66-72 München

COWEN, E.L. (1953): Stress reduction and problem-solving rigidity. J. consultative Psychol. 17

COX, F.N. (1962): Educational streaming and general and test anxiety. Child Develpm. 33, 381-390

DAVIDSON, W.Z., ANDREWS, J.A., ROSS, S. (1956): Effects of stress and anxiety on continuos high speed color naming. J. exp. Psychol. 52, 13-17

DORIS, J. (1959): Test anxiety and blame assignment in grad school children. J. abnorm. soc. Psychol. 58, 181-190

DUNN, J.A. (1968): Anxiety, stress and the performance of complex intellectual tasks: A new look at an old question. J. consultative Clin. Psychol. 32, 669-673

DURNFORD, M.J. (1957): Intellectual performance and classroom behavior of high and low test-anxious children. Ontario, 9, 119-124

ECKSTEIN, H.J. (1958): Kinderneurosen und Schule. Schulwarte 11, 86-93

EYSENCK, H. J. (1957): The dynamics of anxiety and hysteria. London

FELDHUSEN, J.F., KLAUSMEIER, H.J. (1962): Anxiety, intelligence, and achievement in children of low, average, and high intelligence. Child Deficiency 70, 607-608

FRENCH, J.R.P. (1953): Experiments in Field sittings. In: FESTINGER, KATZ: Research methods in the behavioral science, 98-135, New York

FRIEDMAN, M. (1937): The use of ranks to avoid the Assumption of Normality Implicit in the Analysis of Variance. J. Amer. Statist. Ass. 32, 675

FRUHMANN, TH. (1952) Die Überwindung der Schulangst als sozialpädagogische Aufgabe. Pädagogische Provinz 6, 263-270. Frankfurt

FURCK, C.L. (1961): Das pädagogische Problem der Schulleistung. Weinheim

GÄRTNER-HARNACH, V. (1972): Angst und Leistung. Weinheim

GERÉB, G., SZABO, Z. (1963): Ein Verfahren zur psychologischen Untersuchung der Motive der Angst bei Elementarschülern verschiedenen Alters. Praxis der Kinderpsychologie und Kinderpsychiatrie 12, 146-151. Göttingen

GIFFORD, E.M., MARSTON, A.R. (1966): Test anxiety, reading rate, and task experience. J. Educ. Res. 59, 303-306

GLATZ, G.: Vergleichsarbeiten. 3.-9. Schuljahr. Tellus Verlag Essen

GOTTSCHALCH, W., NEUMANN-SCHÖNWETTER, M., SOUKUP, G. (1971): Sozialisationsforschung. Frankfurt

GRAAP, F. (1956): Erscheinungsformen der Angst bei Schulkindern. Praxis der Kinderpsychologie und Kinderpsychiatrie 5, 95-97. Göttingen

GRINKER, R.R., SABSHIN, M.,HAMBURG, D.A., BOARD, F.A., BASOWITZ, H., KORCHIN, S.J., PERSKY, H., CHEVALIER, J.A. (1957): The use of an anxiety-producing interview and its meaning to the Subject. Archives of neurology and psychiatry 77, 406-419

GUILFORD, J.P. (1964): Persönlichkeit. Weinheim.

GUILFORD, J.P. (1965): Fundamental statistics in psychology and education. Mc Graw-Hill

HAECKER, H. (1964): Die Leistung von Volksschülern im schriftlichen Zahlenrechnen bei Übungsarbeiten und Prüfungsarbeiten. Diss. Bonn

HANSELMANN, H. (1950): Examensangst. Schweizerische Erziehungsrundschau 9, 161-180 ST. Gallen

HECKHAUSEN, H., ROELOFSEN, I. (1962): Anfänge und Entwicklung der Leistungsmotivation im Wetteifer des Kleinkindes. Psychol. Forschung 26, 313-397

HECKHAUSEN,H. (1963): Hoffnung und Furcht in der Leistungsmotivation. Meisenheim

HECKHAUSEN, H.(1968): Motivation in der Leistungsgesellschaft. Deutsche Schule 10, 637-648

HECKHAUSEN, H. (1969): Förderung der Lernmotivation und der intellektuellen Tüchtigkeit. In: ROTH, H. (Hrsg.): Begabung und Lernen. Stuttgart

HEBB, D.V. (1955): Drives and the C.N.S. . Psychol. Review 62, dt. in THOMAE, H. (Hrsg.) (1965): Die Motivation menschlichen Handelns. Köln

HELM, J. (1954): Über den Einfluß affektiver Spannungen auf das Denkhandeln. Z. Psychologie 157, 23-105

HENSCHEL, J. (1971): Der Einfluß der Diktiermethode auf die Rechtschreibleistung von Schülern des 9. Schuljahres in schulischen Prüfungssituationen. Unveröff. Examensarbeit zur ersten Prüfung für das Lehramt an Grund- und Hauptschulen. Duisburg

v. HENTIG, H. (1971): Die neuen Ziele und die alte Mentalität. In: LICHTENSTEIN-ROTHER, I. (Hrsg.): Schulleistung und Leistungsschule, 33-43. Bad Heilbrunn

HILGARD, E.R., JONES, L.V., KAPLAN, S.J. (1951): Conditioned discrimination as related to anxiety. J. exp. Psychol. 42, 94-99

HILL, K.T., SARASON, S.B. (1966): The relation of test anxiety and defensiveness to test and school performance over the elementary-school years. Monogr. Society for Research in Child Develpm. 31, 1-76

HOCHHEIMER, W. (1955): Zur Psychologie von Prüfungen. Berlin

HOLFORT, F. (1969): Der Einfluß von Prüfungsdruck auf die Rechenleistung von Hauptschülern. Unveröff. Examensarbeit zur zweiten Prüfung für das Lehramt an Volksschulen. Essen

HOLZKAMP, K. (1972): Kritische Psychologie. Frankfurt

HULL, C.L. (1943): Principles of behavior. New York

INGENKAMP, K.H. (1963): Tests für die Hand des Lehrers. Weinheim

JANKE, W. (1969): Methoden der Induktion der Aktiviertheit. In: SCHÖNPFLUG, W.: Methoden der Aktivierungsforschung, 29-94. Bern

KÖHLER, E. (1936): Entwicklungshilfe als pädagogisches und psychologisches Problem dargestellt am plastischen Gestalten. Weimar

KÖHLER, W. (1940):Dynamics in psychology. New York

KÖHLER, W. (1947): Gestalt psychology. New York

KOFFKA, K. (1935): Principles of gestalt-psychology.

KÜNKEL, F. (1928): Einführung in die Charakterkunde. Leipzig

KVALE, ST. (1972): Prüfung und Herrschaft. Weinheim, Basel

LAZARUS, R.S., DEESE, J., OSLER, S.F. (1952): The effects of psychological stress upon performance. Psychol. Bull. 49, 293-313

LENNERT, R. (1959): Die Prüfung als pädagogisches Problem. Vortrag. Loccum

LEVITT, E.E. (1971): Die Psychologie der Angst. Stuttgart

LEWIN, K. (1926): Vorsatz, Wille und Bedürfnis. Psychol. Forschung 7, 294-385

LEWIN, K. (1927): Gesetz und Experiment in der Psychologie. Symposion 5, Berlin

LEWIN, K. (1931): Die psychologische Situation bei Lohn und Strafe. Leipzig

LEWIN, K. (1935): A dynamic theory of personality. New York

LEWIN, K. (1936): Principles of topological psychology. New York dt.: Grundzüge der topologischen Psychologie Bern (1969)

LEWIN, K. (1938): Experiments on autocratic and democratic atmosphere. The social frontiers 4, 316-319

LEWIN, K. (1939): Field theory and experiment in social psychology: Concepts and methodes. Amer. J. Socio 44, 886-897

LEWIN, K., LIPPITT, R., WHITE, K. (1939): Patterns of aggressive behavior in experimentally created "social climates". J. soc. Psychol. 10, 271-299

LEWIN, K. (1963): Feldtheorie in den Sozialwissenschaften. Bern

LIENERT, G.A. (1962): Verteilungsfreie Methoden in der Biostatistik. Meisenheim

LJUNG, B.O. (1960): Skolmotivation, ängstigkeit och prestation. Lärarhögskolan Stockholm

MANDLER, G., SARASON, S.B. (1952): A study of anxiety and learning. J. abn. soc. Psychol. 47, 166-173

MANN, H.B., WHITNEY, D.R. (1947): On a test of whether one of two random variables is stochastically larger than the other. Amer. Math. Statist. 18, 50

MC COY, N. (1965): Effects of test anxiety on children's performance as a function of instructions and type of task. J. pers. soc. Psychol. 2, 634-641

METZGER, W. (1954): Psychologie. Darmstadt

METZGER, W. (1962): Schöpferische Freiheit. Frankfurt

METZGER, W. (1957): Stimmung und Leistung. Münster

METZGER, W. (1971): Psychologie in der Erziehung. Bochum

MIERKE, R. (1953): Entwicklungsanomalien als Folgeerscheinung seelischer Überforderung. In: Praxis der Kinderpsychologie und Kinderpsychiatrie 2/3, 33-38. Göttingen

MILLER, N.E. (1951): Learnable drives and rewards. In: STEVENS (Ed.): Handbook of experimental psychology, 435-473. New York

MOELLER, M.L. (1969): Psychische Funktion und Wirkungen von Prüfungen. In: SCHÜTZ, M. (Hrsg.): Blickpunkt Hochschuldidaktik 1, 74-80

MORGAN, E., SUTTON-SMITH, B., ROSENBERG, B.G. (1960): Age changes in the relation between anxiety and achievement. Child Develpm. 31, 515-519

NOAK, B. (1972): Der Einfluß der Diktiermethode auf die Rechtschreibleistung von Schülern des 4. Schuljahres in schulischen Prüfungssituationen. Unveröff. Examensarbeit zur ersten Staatsprüfung für das Lehramt an Grund- und Hauptschulen. Duisburg

PAGES, R. (1962): Das Experiment in der Soziologie. In: KÖNIG, R.: Hdb. der empirischen Sozialforschung, 415-450. Stuttgart

PAUL, G.E., ERIKSEN, C.W. (1964): Effects of test anxiety on "real-life" examinations. J. Pers. 32, 480-495

PETERSEN, P., KÖHLER-PETERSEN, E. (1952/1953): Zur Entstehungsgeschichte der pädagogischen Tatsachenforschung. Wiss.Z.Univ. Jena, 119-121

PLATTNER, E. (1968): Die ersten Lebensjahre. Stuttgart

PZ BERLIN (1970): Handbuch der Unterrichtsforschung. Teil I, II, III. Weinheim

RICHTLINIEN und LEHRPLÄNE für die Grundschule - Schulversuch in NRW. Düsseldorf (1969)

GRUNDSÄTZE, RICHTLINIEN und LEHRPLÄNE für die Hauptschule in NRW. Düsseldorf (1968)

SALZMANN, CH. (1966): Pädagogische Überlegungen zum Problem des Berechtigungswesens. Päd. Rsch. 20, 134-151

SANDVOSS, K. (1952): Die Schulangst und ihr Teufelskreis. Niedersächs. Lehrerzeitung 3, Nr. 2

SANDVOSS, K. (1956): Schulangst, ein psychagogisches und heilpädagogisches Problem. Z. für Heilpädagogik 7, 60-68. Nienburg

SARASON, S.B. (1954): The clinical interaction with special reference to the Rorschach. New York

SARASON, S.B., HILL, K.T., ZIMBARDO, P.G. (1964): A longitudinal study of the relation of test anxiety to performance on intelligence and achievement tests. Monogr. Soc. Res. Child Develpm. 29 (Zusammenfassung in: WEINERT: Pädagogische Psychologie. Köln 1970

SARASON, S.B., DAVIDSON, K.S., LIGHTALL, F.F., WAITE, R.R., RUEBUSH, B.K. (1960): Anxiety in elementary school children. New York

SARNOF, J., LIGHTALL, F.F., WAITE, R.R., DAVIDSON, K.S., SARASON, S.B. (1958): A cross-cultural study of anxiety among American and English school children. J. Educ. Psychol. 49, 129-137

SASSENRATH, J.M. (1963): Test anxiety, manifest anxiety, and concept learning without awareness. Psychol. Rep. 12, 71-81

SCHAFER, R. (1954): Psychonalytic interpretation in Rorschach testing. New York

SCHELSKY, H. (1959): Schule und Erziehung in der industriellen Gesellschaft. In: MESSERSCHMIDT, F. (Hrsg.): Weltbild und Erziehung, Bd. 20. Würzburg

SCHWAB, J.R., IVERSON, M.A. (1964): Resistance of high-anxious subjects under ego threat to perception of figural distortion. J.of cons. Psychol. 28, 191-198

SCHWERMER, J. (1966): Die experimentelle Willenspsychologie K. LEWINs. Meisenheim

SELYE, H. (1957): Stress beherrscht unser Leben. Düsseldorf

SIEGEL, S. (1956): Nonparametric Statistics for the behavioral Sciences. Mc Graw-Hill

SILVERSTEIN, A.B. (1966): Anxiety and the quality of human figure drawings. Amer. J. Mental Deficiency 70, 607-608

SINGER, K. (1970): Lernhemmung, Psychoanalyse und Schulpädagogik. München

SOLLE, R. (1969): Dér feldtheoretische Ansatz. In: Hdb. d. Psychol., 7.Bd, Sozialpsychologie, I. Halbband. Gött.

SPENCE, K.W., FARBER, I.E. (1953): Conditioning and extinction as a function of anxiety. J. exp. Psychol. 45, 116-118

SPENCE, K.W. (1958): A Theory of emotionally based drive (D) and its relation to performance in simple learning situations. Amer. Psychol. 13, 131-141

SPENCE, K.W., TAYLOR-SPENCE, J. (1968): The Psychology of Learning and Motivation. New York

STEVENSON, H.W., ODOM, R.D. (1965): The relation of anxiety to children's performance on learning and problem-solving tasks. Child Develpm. 36, 1003-1012

SÜLLWOLD, F. (1959): Bedingungen und Gesetzmäßigkeiten des Problemslöseverhaltens. Referat auf dem 22. Kongreß der Dtsch. Ges. f. Psychologie. Heidelberg

TAUSCH, R., KÖHLER, H., FITTKAU, B. (1966): Variablen und Zusammenhänge der sozialen Interaktion in der Unterrichtung. Z. exp. angew. Psychol. 13, 345-365

TAYLOR, J. (1951): The relationship on anxiety to the conditioned eyelid response. J. exp. Psychol. 41, 81-92

TAYLOR, J. (1953): A personality scale of manifest anxiety. J. abn. soc. Psychol. 48, 285-290

TAYLOR, J. (1956):Level of conditioning and intensity of the adaption stimulus. J. exp. Psychol. 51, 127-131

THURNER, F. (1970): Ängstlichkeit: Eine Persönlichkeitsvariable und ihre Auswirkungen. Psychol. Rsch. 21, 187-213

THURSTONE, L.L. (1928): Attitudes can be measured. Amer. J. Soc. 33, 529-554

TRAMER, M. (1952): Von der Schulangst des Kindes. Schweizer Erziehungsrundschau 9, 161-180

WAGENSCHEIN, M. (1958): Unterricht, Drill oder Erlebnis. In: Das Kind in unserer Zeit. Stuttgart

WALKER, R.E., SPENCE, J.T. (1964): Relationship between digit span and anxiety. J. of consulting psychology 28, 220-223

WEBER, E. (1967): Grundriß der biologischen Statistik. Stuttgart

WERTHEIMER, M. (1957): Produktives Denken. Frankfurt

WESTERMAN HOLSTIJN, A.J. (1966): Verschiedene Definitionen und Auffassungen von "Angst". In: SALZMANN, SCHWIDDER, WESTERMAN HOLSTIJN: Fortschritte der Psychoanalyse. Intern. Jahrbuch zur Weiterbildung der Psychoanalyse Bd. II, 173-188

WIDMER, K. (1955): Gedanken zum Problem der Schulangst. Heilpäd. Werkblätter 24, 194-203. Luzern

WIECZERKOWSKI, W., BASTINE, R., FITTKAU, B., NICKEL, H., TAUSCH, R., TEWES, U. (1969): Verminderung von Angst und Neurotizismus bei Schülern durch positive Bekräftigung von Lehrern im Schulunterricht. Z. Entwicklungspsychol. u. Päd. Psychol. 1, 3-12

WINER, B.J. (1962): Statistical Principles in Experimental Design. Mc Graw-Hill

WINNEFELD, F. (1957): Pädagogischer Kontakt und pädagogisches Feld. Schule und Psychologie 7, Beiheft

YERKES, R.M., DODSON, J.D. (1908): The relation of strength of stimulus to rapidity of habit-formation. J. comp. neurol. Psychol. 18, 459-482

ZIELINSKI, W. (1967): Beziehungen zwischen Ängstlichkeit, schulischer Aktivität, Intelligenz und Schulleistung bei 9-11jährigen Volksschülern. Schule und Psychologie 14, 265-273

ZIMBARDO, P.G., BARNARD, J.W., BERKOWITZ, L. (1963): The role of anxiety and defensiveness in children's verbal behavior.J. Pers. 31, 79-96

NACHTRAG:

DUNCKER, K. (1963): Zur Psychologie des produktiven Denkens. Berlin, Göttingen, Heidelberg

MANDLER, G., COWEN, J. (1958): Test anxiety questionnaires. J. Consult. Psychol. 22, 228-229

A N H A N G

Versuchsanweisungen

4. Schuljahr - Diktat

Erste Phase der Vorbereitung des Hauptversuchs

"Wir wollen heute und auch in den nächsten Tagen einige Übungsdiktate schreiben. Nehmt eure Vorschreib- oder Übungshefte heraus. Wir werden heute so verfahren: Alle Diktate, die wir schreiben, sind nicht vorbereitet. Das Diktat wird einmal ganz vorgelesen, danach jeweils nur ein bestimmter Sinnabschnitt, den ihr auch behalten könnt, dann erst wird der Federhalter aufgenommen und geschrieben (Diktierabschnitte sind in den Vorlagen von GLATZ durch Querstriche / gekennzeichnet). Während der Arbeit werden keine Hilfen gegeben. Wörterbücher werden nicht benutzt. Ich diktiere Satzzeichen mit. Zum Schluß lese ich das Diktat nicht noch einmal vor. Ihr habt aber noch rund zwei Minuten Zeit zum Durchlesen, damit ihr Fehler noch finden könnt. Nach Abschluß des Diktates tauschen wir die Hefte jeweils mit dem Nachbarn aus, der die Fehler kontrolliert."

(Vor der Partnerschaftskontrolle wird die Innentafel geöffnet, auf der der Text des Diktates niedergeschrieben ist.)

Zum Schluß:

"Wie ihr gesehen habt, machen einige von euch noch viele Fehler, obwohl andere sehr erfolgreich gearbeitet haben. Ich werde euch in den nächsten Stunden noch einige Diktate geben."

Zweite Phase der Vorbereitung des Hauptversuches

" Wir werden wieder ein Diktat schreiben. Ihr wißt ja bereits, wie wir das machen. Einmal ganz vorlesen, dann folgen Sinnabschnitte, die ihr niederschreibt, später bekommt ihr zwei Minuten zum Nachsehen. Dann kontrolliert der Nachbar das Ergebnis. Der Text steht hinterher auf der Tafel."

Übungsarbeit

" Wir schreiben wieder ein Diktat ins Übungsheft. Ihr wißt, wie wir das handhaben."

Nach Abschluß des Diktates

" So nun möchte ich auch mal sehen, was ihr nach dieser Übungsserie leisten könnt. Trennt bitte die Seiten mit dem Diktat aus eurem Heft heraus und schreibt euren Namen und Vornamen darauf. Nun sammelt schnell ein!"

Klassenarbeit

" Wir werden heute eine Klassenarbeit schreiben. Ich werde so diktieren, wie ihr es mittlerweile gewohnt seid. Die Arbeit ist wichtig für euch. Deshalb kann ich kein Abschauen gestatten."
(keine weiteren Hinweise)

Prüfungsarbeit

" Wir werden heute eine einmalige Prüfungsarbeit, die für

euch äußerst wichtig ist, schreiben. Wie ihr wißt, werdet ihr die Grundschule am Ende dieses Schuljahres verlassen, um entweder das Gymnasium, die Realschule oder die Hauptschule zu besuchen. Eure Lehrer müssen ganz genau wissen, was ihr könnt, deshalb auch diese wichtige Prüfungsarbeit. In den vergangenen Tagen und Wochen hattet ihr Zeit und die Möglichkeit, euch mit ähnlich schwierigen Diktaten zu befreunden. Strengt euch mächtig an, reißt euch zusammen, damit ihr ein gutes Ergebnis erzielt.

Vorher noch eine Warnung: Demjenigen, der beim Abschauen erwischt wird, wird seine Arbeit sofort abgenommen und mit 'sechs' zensiert.

Bevor ihr anfangt, schreibt als Überschrift 'Prüfungsarbeit', wie es an der Tafel steht und oben links in die Ecke des Blattes euren Vornamen, den Hausnamen und den Namen der Schule."

9. Schuljahr - Rechnen

Erste Phase der Vorbereitung des Hauptversuches

" Im 9. Schuljahr werden viele Stoffgebiete noch einmal aufgearbeitet, die ihr in früheren Schuljahren kennengelernt habt. Dazu gehören auch die Dezimalbruchrechnung und das Rechnen mit gewöhnlichen Brüchen. Diese Gebiete sind so vielfältig, daß man immer wieder üben muß, um sie zu beherrschen. Ich habe auf diesen Blättern Übungsaufgaben zusammengestellt, die ihr in Übungsheften oder auf Bogen, die ihr gerade zur Hand habt, rechnen könnt. Die Aufgaben sind nicht so schwer, daß man auf Nachbarschaftshilfe angewiesen ist. Außerdem wollt ihr ja auch sehen, was ihr bereits könnt. Später werden wir die Aufgaben durchrechnen

und die noch vorhandenen Unklarheiten besprechen. Dann wird der Nachbar nachkontrollieren. Ich werde die Hefte nicht einsammeln, eure Arbeit auch nicht bewerten.

Versucht, die Aufgaben möglichst richtig zu lösen. Die Zeit spielt für die Bewertung eurer Leistung keine Rolle. Es kommt also nicht auf die Zeit an, sondern allein auf die gute Arbeit."
Nach Beendigung der Aufgabenserie:

" So, nun werden wir die Aufgaben der Reihe nach durchgehen und die Ergebnisse vergleichen. Wer noch Schwierigkeiten in der Ausrechnung hat, möge es bitte mitteilen."

Zum Schluß

" Wie ihr gesehen habt, haben manche von euch noch Lücken, obwohl andere fast fehlerfrei gearbeitet haben. Wir werden auch in den nächsten Stunden noch einige Aufgaben rechnen müssen."

Zweite Phase der Vorbereitung des Hauptversuches

" Ich habe euch wieder einige Aufgaben auf diesen Blättern mitgebracht. Sie sind denen, die wir in der letzten Stunde gerechnet haben, ähnlich. Ihr wißt ja nun, wie wir das machen werden. Zeit habt ihr wieder genug. Später kontrolliert der Nachbar wieder die Ergebnisse. Die Aufgaben sind ja nicht so schwer, daß man abschauen muß."

Übungsarbeit

" Die heutigen Übungsaufgaben rechnet ihr wieder wie gewohnt."

(sonstige Hinweise fehlen. Partnerschaftskontrolle wird nicht versprochen. Vorversuche ergaben, daß Rückfragen in dieser Richtung nicht zu erwarten sind.)

Nach Abschluß der Aufgabenserie

"So nun möchte ich auch mal sehen, was ihr nach dieser Übungsserie leisten könnt. Trennt bitte die Seiten mit den gerechneten Aufgaben aus eurem Heft heraus und schreibt euren Namen und Vornamen darauf. Nun sammelt schnell ein!"

Klassenarbeit

"Wir werden heute eine Klassenarbeit schreiben. Es werden wieder Aufgaben gegeben, die euch nicht unbekannt sind. Zeit habt ihr genug. Die Arbeit ist wichtig für euch. Deshalb kann ich kein Abschauen gestatten."
(keine weiteren Hinweise)

Prüfungsarbeit

"Wir werden heute eine einmalige Prüfungsarbeit schreiben, die für eure Zeugnisse sehr wichtig ist. Wie ihr wißt, werdet ihr die Hauptschule am Ende dieses Schuljahres mit Zeugnissen verlassen, von denen es abhängt, ob ihr eine gute Lehrstelle erhaltet oder eine weiterführende Schule besuchen könnt. Eure Lehrer müssen wissen, was ihr könnt, deshalb auch die einmalige Prüfungsarbeit. In den vergangenen Tagen und Wochen hattet ihr Zeit und die Möglichkeit, euch mit den Aufgaben, die ihr in dieser Prüfungsarbeit wiederfindet, zu befreunden.
Strengt euch mächtig an, reißt euch zusammen, damit ihr ein gutes Ergebnis erzielt.

Vorher noch eine Warnung: Demjenigen, der beim Abschreiben erwischt wird, wird seine Arbeit sofort abgenommen und mit 'sechs' zensiert.
Bevor ihr anfangt, schreibt als Überschrift 'Prüfungsarbeit', wie es an der Tafel steht, und oben links in die Ecke des Blattes euren Vornamen, den Hausnamen und den Namen der Schule."

9. Schuljahr - Diktat

Erste Phase der Vorbereitung des Hauptversuches

" In der Schule zeigt es sich immer wieder, daß Rechtschreiben eine wichtige Rolle spielt. Man muß es immer wieder üben, wenn man sich noch verbessern will. Wir wollen heute und auch in den nächsten Tagen einige Übungsdiktate schreiben. Nehmt bitte eure Vorschreib- oder Übungshefte heraus. Wir werden heute so verfahren: Alle Diktate, die wir schreiben, sind nicht vorbereitet. Das Diktat wird einmal vorgelesen, danach jeweils nur ein bestimmter Sinnabschnitt, den ihr auch behalten könnt, dann erst wird der Federhalter aufgenommen und geschrieben (Diktierabschnitte sind in den Vorlagen von GLATZ durch Querstriche / gekennzeichnet). Während der Arbeit werden keine Hilfen gegeben. Wörterbücher werden nicht benutzt. Ich diktiere Satzendzeichen mit. Zum Schluß lese ich das Diktat nicht noch einmal vor. Ihr habt aber noch rund zwei Minuten Zeit zum Durchlesen, damit ihr Fehler noch finden könnt. Nach Abschluß des Diktates tauschen wir die Hefte jeweils mit dem Nachbarn aus, der die Fehler kontrolliert. Ich selbst werde nicht kontrollieren oder gar zensieren."

(Vor der Partnerschaftskontrolle wird die Innentafel geöffnet, auf der der Text des Diktates niedergeschrieben ist.)

Zum Schluß

" Wie ihr gesehen habt, machen einige von euch noch viele Fehler, obwohl andere sehr erfolgreich gearbeitet haben. Ich werde auch in den nächsten Stunden noch einige Diktate geben."

Zweite Phase der Vorbereitung des Hauptversuches

" Wir werden wieder ein Diktat schreiben. Ihr wißt ja bereits, wie wir das machen. Einmal ganz vorlesen, dann folgen Sinnabschnitte, die ihr niederschreibt, später bekommt ihr zwei Minuten zum Nachsehen. Dann kontrolliert der Nachbar das Ergebnis. Der Text steht hinterher auf der Tafel."

Übungsarbeit

" Wir schreiben wieder ein Diktat ins Übungsheft.
Ihr wißt wie wir das handhaben."

(Hinweise auf Partnerschaftskontrolle fehlen.)

Nach Abschluß des Diktates

" So, nun möchte ich auch mal sehen, was ihr nach dieser Übungsserie leisten könnt. Trennt bitte die Seiten mit dem Diktat aus eurem Heft heraus und schreibt euren Namen und Vornamen darauf. Nun sammelt schnell ein!"

Klassenarbeit

" Wir werden heute eine Klassenarbeit schreiben. Ich werde so diktieren, wie ihr es mittlerweile gewohnt seid.

Die Arbeit ist wichtig für euch. Deshalb kann ich kein Abschauen gestatten."

(keine weiteren Hinweise)

Prüfungsarbeit

Wir werden heute eine einmalige Prüfungsarbeit schreiben, die für eure Zeugnisse sehr wichtig ist. Wie ihr wißt, werdet ihr die Hauptschule am Ende des Schuljahres mit Zeugnissen verlassen, von denen es abhängt, ob ihr eine gute Lehrstelle erhaltet oder eine weiterführende Schule besuchen könnt. Eure Lehrer müssen wissen, was ihr könnt, deshalb auch die einmalige Prüfungsarbeit. In den vergangenen Tagen und Wochen hattet ihr Zeit und die Möglichkeit, euch mit den Aufgaben, die ihr in dieser Prüfungsarbeit wiederfindet, zu befreunden.

Strengt euch mächtig an, reißt euch zusammen, damit ihr ein gutes Ergebnis erzielt.

Vorher noch eine Warnung: Demjenigen, der beim Abschauen erwischt wird, wird seine Arbeit sofort abgenommen und mit 'sechs' zensiert.

Bevor ihr anfangt, schreibt als Überschrift 'Prüfungsarbeit', wie es an der Tafel steht, und oben links in die Ecke des Blattes euren Vornamen, den Hausnamen und den Namen der Schule."

Anweisungen für weitere Datenerhebung

Fragebogen

Interesse: Fach (indirekte Befragung)

" Ich habe vor einigen Wochen Schüler des 4. (9.) Schuljahres gefragt, ob sie das Fach Rechnen (Übungsdiktate schreiben) in der Schule besonders gern haben oder nicht. Ich habe eine Menge von Antworten bekommen. Eine Antwort war jeweils besonders häufig dabei. Ihr sollt nun nicht eure Meinung dazu sagen, sondern ihr sollt abschätzen, welche Antworten bei den gefragten Schülern am häufigsten vorkam[en]. Zunächst füllt ihr aber, wenn ihr das Blatt erhalten habt, die Lücken links oben hinter: Name, Vorname, Schule (und evtl. Leistungsgruppe) aus."

" Nun hört bitte ganz genau zu:
Wenn ihr vermutet, daß Schüler in eurem Alter das Fach Rechnen (Übungsdiktate schreiben) in der Schule besonders gern machen, dann kreuzt bitte den linken Kreis über: Schüler mögen Rechnen (schreiben Übungsdiktate gern) an. Vermutet ihr dagegen, daß Schülern in eurem Alter Rechnen (Übungsdiktate) gleichgültig ist (sind) - sie machen es eben nur mit, weil es auf dem Stundenplan steht, sie mögen es weder besonders, noch lehnen sie es ab -, dann kreuzt den mittleren Kreis über: Schülern ist Rechnen (Übungsdiktate schreiben) gleichgültig an. Wenn ihr allerdings meint, daß Schüler in eurem Alter in der Mehrheit das Fach Rechnen (Übungsdiktate schreiben) nicht mögen, dann kreuzt bitte den rechten Kreis an.

Schaut nicht beim Nachbarn nach, denn der kennt das, was ich herausgefunden habe, auch nicht. Er kann nur schätzen. Ich teile euch später mit, was es mit dem Ergebnis auf sich hat."

Knappe Wiederholung

Fragebogen

Interesse, Einstellung: Fach (Schülermeinung)

"Ich habe heute eine Frage für euch, die ihr ohne Aufwand beantworten könnt. Ihr braucht nur etwas anzukreuzen. Ich möchte gern wissen, wer von euch das Fach Rechnen (Übungsdiktate schreiben) in der Schule gern hat und wer nicht. Füllt zunächst, wenn ihr das Blatt erhalten habt, die Lücken oben links hinter Name, Vorname, Schule aus.

Danach hört noch einmal ganz genau zu:

Wenn ihr das Fach Rechnen (Übungsdiktate schreiben) gern mögt, dann kreuzt bitte den Kreis im linken Kästchen an, unter dem steht: Ich mag Rechnen gern. (Ich schreibe Übungsdiktate gern.)

Es kann aber auch sein, daß euch Rechnen (Übungsdiktate schreiben) gleichgültig ist. Ihr macht es eben nur mit, weil es von euch gefordert wird und es im Stundenplan steht. Kreuzt dann bitte den mittleren Kreis über: Mir ist Rechnen (Übungsdiktate schreiben) gleichgültig.

Wenn ihr dagegen Rechnen (Übungsdiktate schreiben) in der Schule nicht mögt - in euch steigt ein Widerwille hoch, wenn das Fach auf dem Stundenplan steht - , dann kreuzt den rechten Kreis über: Ich mag Rechnen nicht (ich schreibe Übungsdiktate nicht gern), an.

Ihr könnt nichts richtig oder falsch machen, denn nur eure Meinung ist gefragt. Deshalb braucht ihr auch beim Nachbarn keine Auskunft einzuholen. Wenn ihr nun noch Fragen habt, dann stellt sie bitte laut.

Knappe Wiederholung.

Fragebogen

Interesse: Fach (Lehrermeinung)

"Es geht in dieser Befragung darum, einigermaßen zuverlässig zu ermitteln, wer von Ihren Schülern Rechnen (Übungsdiktate schreiben) gerne mag und wer nicht. Ich weiß, daß es schwierig ist, von Ihnen ein abgestuftes Urteil über alle Schüler zu verlangen. Sie haben drei Beurteilungsstufen zur Verfügung (+ interessiert / 0 gleichgültig / - Ablehnung). Es soll nicht das Können der Schüler, sondern Ihr diffuser Eindruck aus dem Unterrichtsalltag heraus hinsichtlich des Interesses oder Desinteresses, der Sympathie oder Antipathie, der positiven oder negativen Einstellung gegenüber dem Fach Rechnen (Übungsdiktate schreiben) niedergelegt werden. Ich bitte Sie, Ihren Eindruck unbeeinflußt von der Schülerantwort (es können Urteilsdiskrepanzen auftreten!) per Kreuz in der unteren Zeile niederzulegen."

9. Schuljahr - Rechen- oder Rechtschreibtest

"Ihr habt sicherlich schon mal gehört, daß man mit einem Rechen- (Rechtschreib-)test eure Fähigkeiten auf diesem Gebiet besonders genau einschätzen kann. Ich habe eurem Lehrer versprochen, einen solchen Test mit euch durchzuführen, weil ich auf diesem Gebiet besonders ausgebildet worden bin. Die Ergebnisse werden zur Beurteilung als Bewertungsgrundlage herangezogen."

Anweisungen für die Kontrolluntersuchungen

Nach Einführung des Versuchsleiters und Üben des zu prüfenden Unterrichtsstoffes (wie im Hauptversuch be-

schrieben):

" Wir werden heute den Stoff, den wir in den letzten Stunden durchgearbeitet haben, in drei Übungsarbeiten noch einmal üben. Jedesmal nach einer Arbeit werden die DIN-A-4-Blätter, die ich euch mitgebracht habe, eingesammelt und nachgesehen. Die einzelnen Ergebnisse der Arbeiten werde ich euch allerdings erst nach Abschluß der Serie bekanntgeben. Schaut nicht ab, denn ich möchte sehen, was ihr gelernt habt."

Lebenslauf

Am 29. 4. 1940 wurde ich als Sohn des Zimmerpoliers EMIL HOLFORT und seiner Ehefrau KÄTHE geb. HACKERT in Essen geboren. Meine Schulausbildung schloß ich 1960 mit der Reifeprüfung am Neuspr. Goethe-Gymnasium zu Essen ab. Nach dreijährigem Studium an der Pädagogischen Hochschule Kettwig (Ruhr) - Abschluß: erste Staatsprüfung für das Lehramt an Volksschulen - begann ich im Sommersemester 1963 mit dem Studium der Psychologie, der Pädagogik und Philosophie in Münster. Im Juli 1968 bestand ich die Diplomhauptprüfung für Psychologen an der Universität Münster. Nach anderthalbjähriger Tätigkeit als Lehrer an einer Hauptschule in Essen und erfolgreichem Abschluß der zweiten Staatsprüfung für das Lehramt an Volksschulen wurde ich im Februar 1970 mit der Verwaltung der Stelle eines wissenschaftlichen Assistenten im Seminar für Psychologie an der Gesamthochschule Duisburg (PH Ruhr, Abt. Duisburg) beauftragt.